ATAQUES A REDES MÓVEIS USANDO VULNERABILIDADES DO SS7: UMA ANÁLISE DE TRÁFEGO REAL E PROPOSTA DE AUDITORIA

Luiza Odete Herback de Carvalho Macedo

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Miguel Elias Mitre Campista

Rio de Janeiro
Outubro de 2019

ATAQUES A REDES MÓVEIS USANDO VULNERABILIDADES DO SS7: UMA ANÁLISE DE TRÁFEGO REAL E PROPOSTA DE AUDITORIA

Luiza Odete Herback de Carvalho Macedo

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Miguel Elias Mitre Campista, D.Sc.

Prof. Luís Henrique Maciel Kosmalski Costa, Dr.

Prof. Diogo Menezes Ferrazani Mattos, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
OUTUBRO DE 2019

Macedo, Luiza Odete Herback de Carvalho

Ataques a Redes Móveis usando Vulnerabilidades do SS7: Uma Análise de Tráfego Real e Proposta de Auditoria/Luiza Odete Herback de Carvalho Macedo. – Rio de Janeiro: UFRJ/COPPE, 2019.

XVI, 69 p.: il.; 29, 7cm.

Orientador: Miguel Elias Mitre Campista

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2019.

Referências Bibliográficas: p. 64 – 69.

1. Redes Móveis. 2. Telefonia Celular. 3. SS7. 4. Corrente de Blocos. 5. Ataques. 6. Segurança. I. Campista, Miguel Elias Mitre. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

À minha amada mãe,
Sandra Herback (in memoriam).

Agradecimentos

Agradeço primeiramente a Deus, por ter soberanamente me conduzido até aqui e pela Sua infinita graça e bondade que me acompanham. *Soli Deo gloria.*

Agradeço aos meus pais, Deusdedit e Sandra, que não mediram esforços para me dar uma boa educação. Por todo amor e dedicação empenhados a mim e pelo exemplo de pais, que almejo seguir.

Agradeço ao meu marido, Elienai, por todo apoio que me deu, sendo compreensivo, paciente, amoroso e por ter sido o meu principal incentivador nessa etapa. Por ter me impulsionado a continuar quando eu cansei de correr e por não ter deixado eu desistir, me inspirando sempre a ser uma melhor versão de mim mesma.

Agradeço à minha família, em especial à minha avó Luiza Herbach, à minha tia Sueli Herback, aos meus sogros Edilson e Jussara Macedo e à minha cunhada Elisama Macedo, que foram importantes para a minha caminhada, obrigada por todo apoio necessário. Agradeço aos meus amigos pela força, compreensão e orações.

Agradeço a toda a equipe do Centro de Referência Tecnológica (CRT), em especial ao José Silva, ao Walderson Vidal e Antônio Silvério, pela oportunidade de autodesenvolvimento e por serem muito mais do que gestores, impulsionando-me a ser uma pessoa melhor. Agradeço também ao Roberto Beghini, Fellipe Guimarães e Mauro Rodrigues pela disponibilização dos dados e suporte. Também agradeço aos Responsáveis Técnicos do CRT, principalmente à Carolina Neves, Laila Sousa e Zeneide Veras, pela amizade, companheirismo e compreensão. Agradeço ao amigo Leon Porto, por ter me ajudado quando surgiam dúvidas, sendo sempre prestativo.

Gostaria de agradecer ao meu orientador, Miguel Campista, pela confiança em mim depositada, pela compreensão nos períodos mais difíceis desta jornada, pela paciência e por ter me ajudado significativamente para que este trabalho fosse desenvolvido. Também agradeço ao Grupo de Teleinformática e Automação (GTA), pela oportunidade de aprendizado, receptividade e ajuda necessária, em especial à minha colega Mariana Maciel.

Agradeço também aos professores Luís Henrique e Diogo Ferrazani, pela participação na banca examinadora desta dissertação e aos funcionários do Programa de Engenharia Elétrica da COPPE/UFRJ, pela presteza no atendimento na secretaria do Programa.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

ATAQUES A REDES MÓVEIS USANDO VULNERABILIDADES DO SS7: UMA ANÁLISE DE TRÁFEGO REAL E PROPOSTA DE AUDITORIA

Luiza Odete Herback de Carvalho Macedo

Outubro/2019

Orientador: Miguel Elias Mitre Campista

Programa: Engenharia Elétrica

O Sistema de Sinalização nº 7 (SS7) define uma pilha de protocolos usados principalmente na troca de sinalização das redes de Provedores de Serviço, como por exemplo, em redes móveis. Originalmente, tais protocolos foram baseados em relações de confiança mútua entre as partes, sem preocupação com segurança de rede. Com o surgimento do chamado "mundo IP"e o crescimento do número de operadoras, as redes móveis ficaram expostas a ataques em SS7. Mediante o uso do SS7, atacantes podem localizar usuários, obter dados privados e até provocar a indisponibilidade de serviços. Várias contramedidas foram propostas para os ataques em SS7, como *firewalls* e criptografia assimétrica, mas nenhuma delas foi totalmente efetiva. A partir da análise de dados de um tráfego real, este trabalho possui como objetivos: a avaliação da vulnerabilidade da rede de uma grande operadora brasileira de telecomunicações e a caracterização das ameaças obtidas para modelagem dos atacantes. Adicionalmente, é feita a proposta do uso da tecnologia de corrente de blocos como forma de introduzir auditabilidade e rastreabilidade das operações de rede, servindo como complemento às contramedidas já existentes. Desse modo, torna-se possível identificar ameaças e determinar o impacto das mesmas na rede, culminando na melhoria da segurança das redes móveis. Por fim, é feita a avaliação da viabilidade da proposta, através de medições do consumo de recursos computacionais, vazão e latência das transações da corrente de blocos, a fim de mensurar o resultado da adoção da tecnologia aos elementos de rede já implantados atualmente. Através da análise de desempenho, mostra-se que é possível implantar a tecnologia proposta, sem causar grandes impactos a infraestrutura existente na rede da operadora.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

MOBILE NETWORK ATTACKS USING SS7 VULNERABILITIES: A REAL TRAFFIC ANALYSIS AND AUDIT PROPOSAL

Luiza Odete Herback de Carvalho Macedo

October/2019

Advisor: Miguel Elias Mitre Campista

Department: Electrical Engineering

Signaling System 7 (SS7) defines a stack of protocols used primarily in the interconnection of networks from Service Providers, for example, in mobile networks. Originally, such protocols were based on mutual trust relationships between components, i.e., SS7 was not designed focusing on network security. With the emergence of the so-called "IP world" and the growth in the number of carriers, mobile networks got exposed to attacks on SS7. By means of SS7, attackers can locate users, obtain private data, and even trigger a denial of service attack. Several countermeasures have been proposed for SS7 attacks, such as firewalls and asymmetric cryptography, but none of these were fully effective. From the data analysis of a real traffic, this work has as purposes: the network vulnerability evaluation of a large brazilian telecommunications operator and the characterization of the threats obtained for modeling the attackers. Additionally, it is proposed to use blockchain technology as an approach to introduce auditability and traceability of network operations, suited as a complement to existing countermeasures. This makes it becomes possible to identify threats and determine their impact on the network, leading to an improved security of mobile networks. Finally, the proposal feasibility is evaluated through computational resource consumption measurements, flow and latency of blockchain transactions, in order to measure the result of the adoption of the technology to the network elements already implemented. Through performance analysis, it is shown that it is possible to adhere to the proposed technology, without causing major impacts to the existing infrastructure in the operator's network.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas

3GPP	*3rd Generation Partnership Project*
ATI	*anyTimeInterrogation*
AuC	*Authentication Center*
BCCH	*Broadcast Common Control Channel*
BSC	*Base Station Controller*
BSS	*Base Station Subsystem*
BSSMAP	*Base Station Subsystem Mobile Application Part*
BTS	*Base Transceiver Station*
CAMEL	*Customizable Applications for Mobile Enhanced Logic*
CAPEX	CAPital EXpenditure
CdPA	*Called Party Address*
CFT	*Crash Fault Tolerance*
CgPA	*Calling Party Address*
CSFB	*Circuit Switch Fallback*
DTAP	*Direct Transfer Application Part*
EIR	*Equipament Identity Register*
FDMA	*Frequency Division Multiple Access*
GGSN	*Gateway GPRS Support Node*
GMSC	*Gateway Mobile Switching Center*
GPRS	*General Packet Radio Service*

GSM *Global System for Mobile Communications*

GSMA *Global System for Mobile Communications Association*

GT *Global Title*

HC *Hyperledger Caliper*

HLF *Hyperledger Fabric*

HLR *Home Location Register*

HNB *Home NodeB*

HSS *Home Subscriber Server*

IMEI *International Mobile Equipment Identity*

IMSI *International Mobile Subscriber Identity*

IMS *IP Multimedia Subsystem*

IP *Internet Protocol*

LTE *Long Term Evolution*

M2M *Machine-to-Machine*

MAP *Mobile Application Part*

MCC *Mobile Country Code*

MNC *Mobile Network Code*

MS *Mobile Station*

MSC *Mobile Switching Center*

MSISDN *Mobile Station International Subscriber Directory Number*

MSIN *Mobile Subscriber Identity Number*

MVNO *Mobile Virtual Network Operators*

NSS *Network Switching Subsystem*

OSI *Open Systems Interconnection*

PBFT *Practical Byzantine Fault Tolerance*

PCH *Paging Channel*

PoA *Proof-of-Authority*

PoS *Point of Sale*

PoW *Proof-of-Work*

PRD *Permanent Reference Document*

PSI *provideSubscriberInfo*

RNC *Radio Network Controller*

SCP *Service Control Point*

SDR *Software Defined Radio*

SGSN *Serving GPRS Support Node*

SIGTRAN *Signaling Transport*

SMPP *Short Message Peer-to-Peer*

SMS *Short Message Services*

SS7 *Signalling System Nº 7*

SSP *Service Switching Point*

STP *Signal Transfer Point*

TCAP *Transaction Capability Application Part*

TDMA *Time Division Multiple Access*

UMTS *Universal Mobile Telecommunication System*

UMTS *Universal Mobile Telecommunications System*

USRP *Universal Software Radio Peripheral*

UTRAN *Universal Terrestrial Radio Access Network*

VLR *Visitor Location Register*

VoLTE *Voice Over Long Term Evolution*

WCDMA *Wide-Band Code-Division Multiple Access*

Capítulo 1

Introdução

As redes de telecomunicações móveis revolucionaram a forma de comunicação humana e trouxeram implicações nas relações interpessoais, influenciando na identidade do indivíduo bem como seu comportamento e cosmovisão. A facilidade de acesso à *Internet*, a disseminação da informação em tempo real, o surgimento das redes sociais, a acessibilidade a novos serviços e a mobilidade de voz, são alguns dos fatores que contribuem na formação da sociedade pós-moderna em aspectos culturais, políticos e econômicos [1].

A evolução das tecnologias sem fio é contínua e permanecerá abrindo novos precedentes para aplicações e meios de comunicação que antes eram inconcebíveis, como por exemplo, o novo paradigma da Internet das Coisas (*Internet of Things* - IoT) onde objetos inteligentes podem comunicar-se [2]; e a Internet Tátil, na qual várias aplicações de realidade virtual com latência de 1 ms são desenvolvidas para melhorar a experiência do usuário [3].

Segundo o relatório "*The Mobile Economy 2019*", elaborado pelo *Global System for Mobile Communications Association* (GSMA), no fim de 2018, as redes móveis chegaram a 5,1 bilhões de assinantes, o que representa 67% da população mundial [4]. Obviamente, a necessidade constante de maior largura de banda, a exigência de mínima latência e o aumento da qualidade de experiência do usuário promovem a aceleração do desenvolvimento de novas tecnologias como suporte às novas aplicações, como a quinta geração de redes móveis, o 5G. Isto impõe aos provedores de serviço uma rápida adequação de infraestrutura a fim de não perder oportunidades de negócio e valor agregado, e reacende a preocupação com segurança de rede, tendo em vista a elevada quantidade de conexões.

Apesar da grande evolução tecnológica, a segurança das redes móveis permanece comprometida, ainda nas gerações móveis mais atuais. Uma das razões é o uso legado do SS7 ou Sistema de Sinalização por Canal Comum n^{o}7, que é uma antiga pilha de protocolos de sinalização usada principalmente na troca de sinalização de redes móveis entre operadoras, quando há necessidade de expansão da cobertura de

rede. Quando os protocolos da pilha SS7 foram concebidos, a partir dos anos 80, não se pensava em segurança como hoje, já que as relações entre os provedores de serviço eram mutuamente confiáveis [5]. Porém, em 2008, durante a "*25th Chaos Computer Club Conference*", os principais ataques em SS7 foram expostos, anteriormente desconhecidos [6]. Dentre os ataques possíveis estão: fraudes, obtenção de dados de clientes, localização de usuários, uso indevido da rede da operadora para a geração de chamadas massivas, terceirização de serviços e até indisponibilidade da rede ou de serviços. O reflexo do relaxamento da segurança no projeto do SS7 pode ser mais percebido atualmente, a partir do mundo baseado em IP (*Internet Protocol*) e seus frequentes maus usos da rede.

Dada a criticidade do problema de segurança em redes móveis, as organizações de padronização para telecomunicações, tais como o *3rd Generation Partnership Project* (3GPP) e o GSMA, preocuparam-se em estabelecer novos padrões de segurança para as gerações móveis mais recentes. Vários requisitos de segurança para interconexão de redes de provedores de serviços foram incluídos em documentos denominados *"Permanent Reference Document"* (PRD). Dentre esses documentos estão o IR.70 [7], que descreve ataques de fraude e interceptação SMS em sinalização SS7; o IR.77 [8], definindo características de interconexão de *Backbones*; e o IR.88, que apresenta regras de segurança em situações de *roaming* e de *Circuit Switch Fallback* (CSFB) para redes *Long Term Evolution* (LTE) [9]. Embora as medidas de regulamentação tenham sido tomadas, o 3GPP assume que os padrões são estabelecidos entre nós confiáveis e entre redes dentro do domínio de confiança da operadora, ou seja, com quem ela possui contratos. Na prática, todavia, essa realidade não se reflete totalmente. Além disso, ataques como o de rastreamento são difíceis de mitigar, visto que os telefones móveis enviam sinais de rádio em *broadcast* de forma periódica para realizar o *"attach"* (ou registro) na rede da operadora. Isso quer dizer que, se o atacante identificar a célula de registro, é possível rastrear a localização do usuário.

Como solução definitiva, pode-se sugerir a completa substituição das redes móveis legadas (2G e 3G) que utilizam o sistema SS7. Esta solução, entretanto, está longe de ser alcançada. No último relatório emitido pelo GSMA que apresenta a situação atual das redes móveis no ano de 2019, mostrou-se que somente em 2018 o 4G ultrapassou o uso do 2G, tornando-se a tecnologia líder das gerações móveis, com 3,4 bilhões de usuários, compreendendo 47% da totalidade de assinantes a nível mundial [4], enquanto o 2G ainda possui 36% da quantidade total de usuários. Ainda de acordo com o relatório, a expectativa é que em 2025, o 4G seja responsável por 59% das conexões móveis, seguido pelo 3G com 20%, 5G com 15% e finalmente o 2G, com 5%. Independentemente do declínio do 2G, a manutenção das redes com esta tecnologia ainda se faz necessária para suporte a aplicações corporativas

M2M (*Machine-to-Machine*) legadas, utilizadas por exemplo, em POS (*Point of Sale*), comumente conhecidas como máquinas de cartão de crédito. Portanto, o 2G permanece sem data definida de descontinuidade e perspectiva de desligamento por parte das operadoras.

A partir do 4G, o SS7 foi substituído pelo protocolo Diameter. O seu sucessor, porém, também possui vulnerabilidades, sendo tão comprometido em segurança quanto o SS7 para o ataque de rastreamento de usuário, por exemplo [10–12]. Ademais, a conversão do SS7 para Diameter tem sido feita de forma gradual a fim de manter serviços legados, isto é, os nós de interconexão de centrais de rede permanecem híbridos, suportando SS7 e Diameter. Outro fator é que o LTE não suporta voz nativamente já que é uma rede puramente baseada em pacotes de dados. Logo, se a operadora não possuir um *Core* IMS (*IP Multimedia Subsystem*), que dá suporte a chamadas VoLTE (*Voice Over Long Term Evolution*), é necessário que haja um rebaixamento de tecnologia (ou CSFB) para 3G ou 2G para disponibilização dos serviços de voz [13]. Essas características apontam que mesmo as redes mais recentes de telefonia móvel não estão imunes aos ataques em SS7.

O presente trabalho possui como objetivos: avaliação da exposição de rede de uma grande provedora de serviços de telecomunicações do Brasil ao problema de vulnerabilidades em SS7; a caracterização das ameaças encontradas para modelagem dos atacantes; a proposta do uso da tecnologia de corrente de blocos como forma de introduzir auditabilidade e rastreabilidade das operações de rede e a análise da viabilidade do modelo de corrente de blocos proposto em um ambiente real.

Para a verificação da vulnerabilidade da rede, foi feita uma coleta de dados do tráfego real de roaming internacional da operadora durante o período de 5 meses. A partir da análise dos dados, são apresentados resultados mostrando as principais ameaças encontradas na rede da operadora, a distribuição da origem e destino das ameaças, a duração das mesmas e a periodicidade das ocorrências. Observa-se que as ameaças possuem diferentes objetivos, sendo que os casos de realocação rápida (*Fast Relocation*), de cancelamento de um usuário (*Cancel Completed*) e de obtenção de dados (*"Send Routing Information for Short Messages"* e *"Provider-SubscriberInfo Completed"*) são os mais disparados, respondendo por mais de 89% de todas as ameaças sofridas pela rede da operadora. Estas ameaças podem ser classificadas como pertencentes às categorias de Interceptação (onde uma ligação ou uma mensagem SMS podem ser capturadas e/ou redirecionadas para o atacante), Fraude (uso ilegal da rede ou de serviços, atribuindo custos à outra parte) e Negação de Serviço (caso em que os serviços tornam-se indisponíveis a um assinante ou um elemento de rede fica inoperante). Os resultados ainda mostram que o intervalo entre ameaças consecutivas apresentam distribuição exponencial, sendo que muitas ocorrem de forma consecutiva ou ao mesmo tempo.

A partir da identificação do problema e evidências de vulnerabilidades na rede da operadora, a introdução da tecnologia *Blockchain*, ou corrente de blocos, é proposta, a fim de promover a auditoria das movimentações da rede que possibilitem o mapeamento de rede móvel e o desenvolvimento de estratégias que mitiguem as ameaças, podendo ser utilizada como complemento às contramedidas já existentes.

As características inerentes à corrente de blocos como transparência, imutabilidade e privacidade têm despertado interesse na resolução de problemas em redes móveis. Babu *et al.* propuseram o uso da corrente de blocos como meio de inserção de segurança na rede SS7 [14]. Os atacantes aproveitam-se de uma limitação de segurança no canal de comunicação entre o usuário e a estação base, e escutam a transmissão da interface aérea, descobrindo informações de localização do usuário que são divulgadas via *broadcast*. Ao invés desta difusão, os elementos de rede seriam inseridos na corrente de blocos e utilizariam criptografia com chaves privadas e públicas para a comunicação, onde somente a chave pública é propagada. Como todas as estações base são nós pertencentes a uma corrente de blocos comum, não há necessidade de autenticação do usuário cada vez que este mude de BTS, pois tais informações públicas já são conhecidas por todos os nós através da replicação das informações dentro da cadeia de blocos. Mafakheri *et al.* propuseram o uso da corrente de blocos para prover uma forma segura de autenticação e armazenamento de informações de usuários em redes 4G, através da descentralização do banco de dados HSS (*Home Subscriber Server*) [15]. Na proposta, os processos de registro e desativação de usuários são feitos através de um contrato inteligente via corrente de blocos ao invés do procedimento normal utilizando o HSS. A ideia é descentralizar as informações dos usuários para que não fiquem dependentes de um único elemento de rede, que é susceptível a falhas. No entanto, o esquema proposto concentra-se na possibilidade de falha ou vulnerabilidade apenas do banco de dados da rede, não se preocupando com requisições legítimas originadas por usuários maliciosos que dispariariam ataques ainda que em um ambiente descentralizado.

A proposta deste trabalho não se baseia em apenas um único tipo de ameaça e também não possui como objetivo a mitigação de ataques. Entende-se que para o desenvolvimento de uma solução efetiva de mitigação e como forma de implementação de um ciclo de segurança, o primeiro passo é a auditoria da rede e o conhecimento das mensagens trocadas na mesma, possibilitando a verificação do cenário de vulnerabilidade. Esta avaliação pode ser obtida de maneira confiável através de uma auditoria por meio da corrente de blocos.

Por fim, para meios de avaliação da viabilidade de implantação da corrente de blocos ao ambiente de redes móveis, foi desenvolvido um contrato inteligente que implementa os principais pontos da proposta na plataforma *Hyperledger Fabric*, que é mantida pela *Linux Foundation* e permite que aplicações específicas de correntes

de blocos privadas e permissionadas sejam desenvolvidas. Para realização dos testes de desempenho e verificação da factibilidade do uso da corrente de blocos, o contrato inteligente foi executado no *Hyperledger Caliper*, que é uma plataforma para aferição de desempenho da corrente de blocos. Através desta ferramenta, alterando-se a quantidade de transações emitidas e o tamanho do bloco de acordo com trabalhos anteriores [16–18], foi possível emitir resultados de vazão, latência e recursos computacionais, como o consumo de memória e CPU.

A presente dissertação está organizada da seguinte forma. O Capítulo 2 fornece a fundamentação teórica sobre redes móveis e também introduz a pilha de protocolos SS7, bem como o seu funcionamento na interconexão de redes, principais vulnerabilidades e contramedidas. O Capítulo 3 apresenta a análise dos dados coletados bem como a discussão dos resultados obtidos pela caracterização. O Capítulo 4 apresenta a teoria sobre a tecnologia de corrente de blocos bem como a proposta de auditoria para redes móveis. Também são apresentados neste capítulo os resultados experimentais da análise de desempenho da proposta. O Capítulo 5 apresenta os trabalhos relacionados. Por fim, o Capítulo 6 conclui o trabalho e apresenta desafios futuros.

Capítulo 2

Redes Móveis e o Sistema SS7

Nas próximas seções deste capítulo, será mostrado um breve histórico das redes móveis legadas e suas arquiteturas. Além disso, será revisto o contexto do sistema de sinalização por canal comum n^{o}7, bem como o seu papel dentro da interconexão de redes móveis. Também será abordado o problema de segurança advindo deste sistema, mostrando os principais tipos de ataques e contramedidas.

2.1 Evolução e Arquitetura das Redes Móveis

A arquitetura tradicional das tecnologias 2G ou GSM (*Global System for Mobile Communications*), 3G ou UMTS (*Universal Mobile Telecommunication System*) e 4G ou LTE é composta por uma rede de acesso e uma rede de núcleo. Neste trabalho, o foco será dado apenas nas duas primeiras gerações de redes móveis, portanto a arquitetura do 4G não será apresentada. Ademais, como a interconexão com redes externas é comum, por exemplo, a interconexão com redes IP ou redes de telefonia fixa, considera-se neste trabalho as redes externas ou de interconexão como sendo um subsistema adicional à arquitetura por questões didáticas.

A Figura 2.1 mostra a interconexão entre as redes de acesso e de núcleo das arquiteturas 2G/3G, assim como a interconexão com as redes externas. Os elementos inseridos no retângulo azul pertencem a ambas arquiteturas, compondo o núcleo legado da rede, que utiliza comutação por circuitos. Já o retângulo amarelo representa os elementos que surgiram posteriormente para encaminhamento de pacotes de dados. Esses elementos foram acoplados ao 2G permitindo esta nova maneira de comutação. As linhas contínuas representam a ligação física do nó da rede e as linhas pontilhadas mostram conexões lógicas. Cada subsistema é apresentado em maiores detalhes nas subseções a seguir.

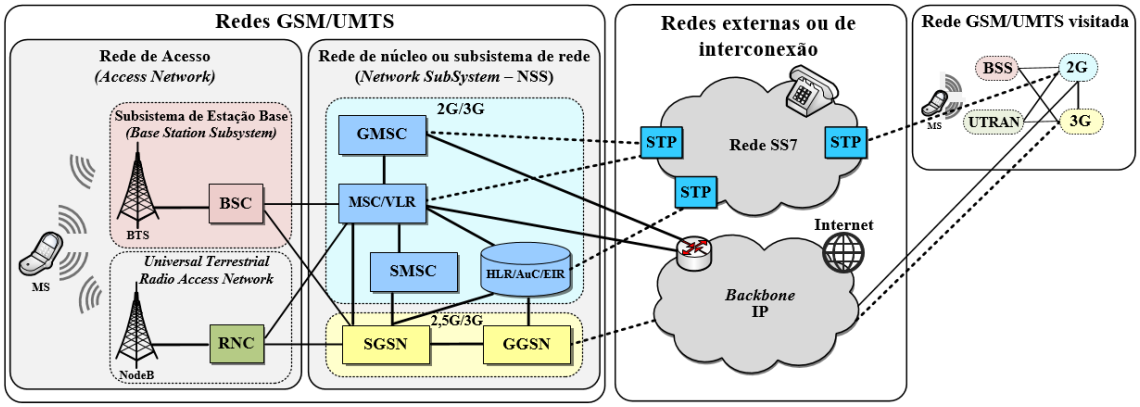


Figura 2.1: Topologia básica da arquitetura da rede móvel 2G (GSM) e 3G (UMTS) e a sua interconexão com redes externas.

2.2 Arquiteturas de Rede 2G e 3G

A rede de acesso, comumente conhecida como Subsistema de Estação Base (*Base Station Subsystem - BSS*) ou simplesmente Rede de Acesso Rádio (*Universal Terrestrial Radio Access Network - UTRAN*), respectivamente às tecnologias 2G e 3G, é o subsistema que viabiliza a entrada do usuário na rede, provendo a conexão do terminal móvel com os elementos de rede responsáveis pela disponibilização dos serviços de voz e dados. A rede de acesso é composta por Estações Transceptoras Base (BTS -*Base Transceiver Station*) e NodeB, para redes 2G e 3G, respectivamente e pelo controlador (BSC - *Base Station Controller* e RNC - *Radio Network Controller*). As Estações Transceptoras Base possibilitam a comunicação entre a Estação Móvel (*Mobile Station* - MS) e a rede. As BTS's são compostas por antenas, transmissores, receptores e processadores de sinais e proveem serviços para uma área de cobertura chamada célula. A principal melhoria entre a BTS e sua evolução, a NodeB, se dá pelo método de acesso, de FDMA/TDMA (*Frequency Division Multiple Access/ Time Division Multiple Access*) para a técnica de codificação WCDMA (*Wide-Band Code-Division Multiple Access*), onde os usuários utilizam um código único ao invés de transmissão por diferentes frequências ou *slots* de tempo. Isto permitiu uma otimização no sistema, alcançando maiores larguras de banda. O controlador de estação base (BSC/RNC) é o elemento central da rede de acesso, sendo responsável pela alocação de recursos de rádio para as BTS's e pelo monitoramento da qualidade da potência emitida a fim de garantir a manutenção das conexões de rádio e a possibilidade de *handover* quando necessário. A principal diferença entre as duas gerações de controladores foi a centralização das funcionalidades de controle apenas na RNC.

A rede de núcleo ou o subsistema de comutação de redes NSS (*Network Switching Subsystem*) é responsável por todas as funções de controle da rede tais como operação

e manutenção, qualidade de serviços e realização de chamadas. Arquiteturalmente, o NSS se subdivide ainda em três componentes principais: central de comutação, bases de dados de usuários e elementos de interconexão.

A central de comutação (*Mobile Switching Center* - MSC) é o principal elemento da rede de núcleo. A MSC é responsável pelo controle de chamadas, funções de controle das redes de acesso, interconexão de redes, faturamento (*billing*), localização dos usuários e *handover*. Além disso, a central de comutação faz também a interface de sinalização entre as redes de acesso e redes externas, tais como: redes de comutação por circuitos, redes SS7 e redes baseadas em IP, vide Figura 2.1.

As centrais de comutação que não possuem interconexão com redes de acesso são conhecidas por centrais de comutação *gateway* (*Gateway Mobile Switching Center* - GMSC). Tais centrais realizam a interconexão entre centrais e atuam como ponto de acesso para outras redes, sendo estas móveis ou fixas. Em situações de *roaming*, o usuário se comunica com sua rede de origem (*home*) através da GMSC.

A definição de rede de origem está relacionada à mobilidade do usuário. Cada vez que o assinante liga o seu aparelho celular, é feita uma tentativa de conexão à rede. Isso requer que os usuários da rede estejam previamente registrados para que a rede possa identificá-los e então disponibilizar os serviços que estão habilitados para o usuário. Para isso, a rede possui elementos com função de banco de dados para armazenamento de informações dos usuários. Como seria inviável ter uma réplica destas informações em todos os bancos de dados da rede móvel, foi criado o conceito de rede de origem e rede visitada. Quando o usuário sair da cobertura de sua rede de origem (ou da sua MSC), a rede visitada consultará a mesma para copiar os dados deste usuário temporariamente. Os responsáveis por estas funções são o HLR (*Home Location Register*) e o VLR (*Visitor Location Register*).

O HLR é o banco de dados de origem do assinante, no qual todos os dados permanentes dos usuários são armazenados. Dentre as informações armazenadas no HLR estão: o número do telefone do assinante (*Mobile Station International Subscriber Directory Number* - MSISDN); o registro do seu SIM Card (ou *International Mobile Subscriber Identity* - IMSI), que é a identificação do usuário na rede; o tipo de plano contratado e quais serviços suplementares o usuário tem acesso. O HLR é usado na ativação e desativação destes serviços fornecidos. Para as tecnologias 3G e 4G, o HLR é substituído pelo HSS (*Home Subscriber Server*), que possui as mesmas funcionalidades do seu antecessor e inclusive compartilham o mesmo *hardware*. Porém, no 4G, utiliza-se o protocolo Diameter para troca de mensagens na rede.

O banco de dados temporário é o VLR, que compartilha o mesmo *hardware* da central (MSC) e atua principalmente quando existe uma migração do usuário da rede de origem para a rede visitada. Na situação de *roaming*, por exemplo, o VLR da rede visitada consulta os dados do HLR do usuário, faz uma cópia e o atribui a

si mesmo, atualizando a localização atual do assinante. O VLR local armazena esta informação do posicionamento do usuário, que é atualizada cada vez que o mesmo muda de área. Com isso, é feito o registro do usuário na rede visitada, o qual pode continuar usufruindo dos serviços de sua operadora.

Além do HLR e do VLR, há ainda alguns bancos de dados com serviços especiais. O EIR (*Equipment Identity Register*) é um destes, muito usado no bloqueio de aparelhos roubados. O EIR permite o cadastro dos números seriais (ou *International Mobile Equipment Identity* - IMEI) dos aparelhos móveis a fim de liberar ou restringir o uso da rede filtrando pelo IMEI. Finalmente, tem-se o AuC (*Authentication Center*), que é geralmente associado a um dado HLR e tem como função realizar a autenticação de usuários na rede. Cada *SIM Card* possui uma chave de autenticação do usuário (Ki) que, ao combinar-se com o IMSI do usuário e um número aleatório, formam um valor que deve ser comparado com outro armazenado na rede (por meio de chave simétrica). Se forem compatíveis, os serviços são liberados ao usuário.

Como o GSM não suporta pacotes de dados, as redes GPRS (*General Packet Radio Service*) foram desenvolvidas pelo ETSI com esta funcionalidade e incorporadas ao sistema, que tornou-se o 2,5G. Portanto, na arquitetura GSM, incluem-se dois elementos GPRS, que se assemelham às centrais MSC e GMSC, porém com suporte ao protocolo IP, são eles: o SGSN e o GGSN. O SGSN (*Serving GPRS Support Node*) é conectado à rede de acesso e executa funções similares às da MSC, incluindo localização de usuários, registro, controle de acesso de usuários, ativação e desativação de serviços. Já o GGSN (*Gateway GPRS Support Node*) é o espelho da central GMSC e é o ponto de interconexão entre redes IP, como o IMS (*IP Multimedia Subsystem*) e o GSM. Este é conectado ao SGSN via *backbone* IP.

2.3 Sistema de Sinalização por Canal Comum n^{o}7

O propósito da sinalização é estabelecer uma linguagem comum, com sintaxe e parâmetros padronizados para criação de um diálogo, a fim de que duas centrais telefônicas se comuniquem.

No método de sinalização tradicional, as informações de controle da rede e de conversação compartilham os mesmos circuitos. Já a sinalização por canal comum permite que as informações de sinalização e controle trafeguem por um único canal de dados dedicado. Isso aumenta a eficiência da rede, pois os canais de voz ficam ocupados apenas pelo tráfego da conversação. O Sistema de Sinalização por Canal Comum n^{o}7 comumente conhecido como SS7 foi padronizado pela ITU-T, tendo a sua primeira publicação em 1980.

O SS7 possui 3 tipos de nós em sua arquitetura, que são chamados de Pontos de Sinalização (*Signaling Points - SPs)* [5], são eles:

- *Service Switching Point* (SSP), que processa o tráfego de voz e pode originar ou receber mensagens de sinalização, correspondendo às centrais telefônicas;
- *Service Control Point* (SCP) é uma base de dados que auxilia no estabelecimento de uma nova conexão, funcionando como uma espécie de memória *cache*;
- *Service Transfer Point* (STP) é responsável por encaminhar mensagens de sinalização SS7 entre os demais nós da arquitetura SS7 (SSPs e SCPs), executando papel similar ao de um roteador IP.

Na Figura 2.2, tem-se um exemplo de uma arquitetura SS7 com pontos de sinalização de diferentes funções. Ela se interliga com a topologia de redes móveis descrita na seção anterior, pois o SS7 é usado no 2G e no 3G para entrega de chamadas, fornecimento de serviços suplementares, gerenciamento de mobilidade, *roaming*, SMS (*Short Message Service*), serviços pré-pago e e autenticação do usuário [5].

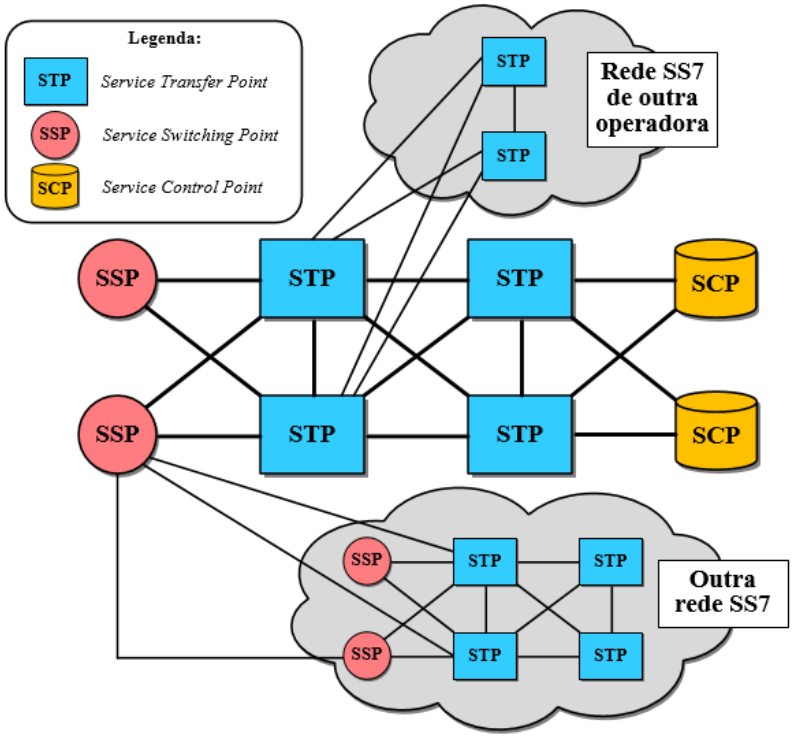


Figura 2.2: Topologia de referência de uma rede SS7.

O SS7 é composto de dois subsistemas de protocolos: o Subsistema de Transferência de mensagens (que preocupa-se com o transporte das informações) e o Subsistema de Usuário (responsáveis pelos serviços e aplicações). Cada subsistema é dividido por funções, formando uma pilha de camadas de protocolos, que funcionalmente pode ser relacionada com a pilha do modelo *Open System Interconnection* (OSI). Por exemplo, a camada de transferência de mensagens (MTP - *Message*

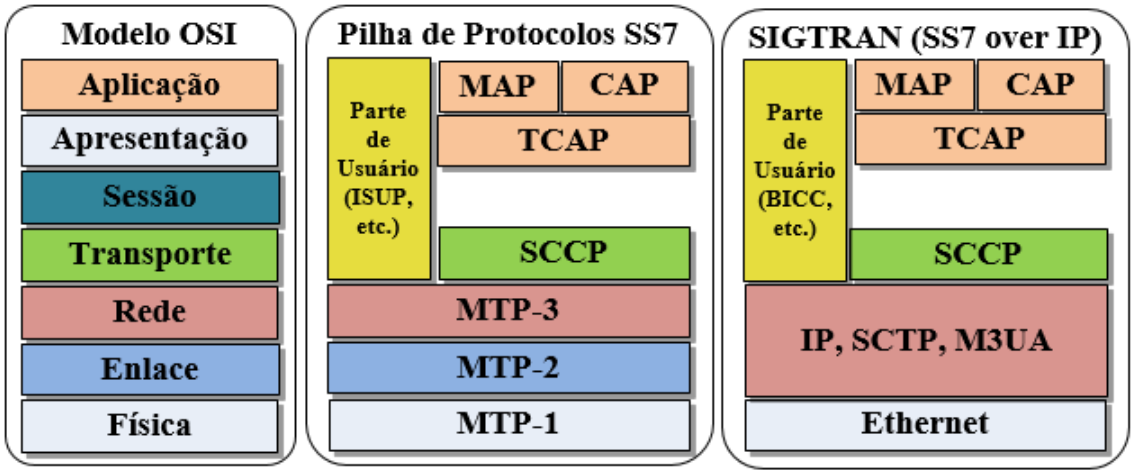


Figura 2.3: Comparação entre as pilhas de protocolos OSI, SS7 e SIGTRAN.

Transfer Part) é equivalente às camadas física (MTP-1), de enlace (MTP-2) e de rede (MTP-3) do modelo OSI. Já os protocolos da parte de usuário e adicionais à arquitetura, equivalem a camada de aplicação do modelo OSI.

Com o surgimento da Internet, porém, tornou-se necessária a interoperabilidade entre a rede SS7 e as redes IP. Para tanto, em 1999, o IETF criou o SIGTRAN (*Signaling Transport*), que tem como objetivo promover a integração da rede de circuitos comutados para rede de pacotes de dados. A arquitetura SIGTRAN define um conjunto de protocolos que encapsulam mensagens nativas SS7 em datagramas IP [19]. A comparação entre o SIGTRAN, o SS7 e o modelo OSI está explicitado na Figura 2.3. Nota-se que com o SIGTRAN, também é possível a uma aplicação SS7 estar baseada diretamente em IP.

O MAP (*Mobile Application Part*), exibido na Figura 2.3, foi padronizado pelo 3GPP em 1999 [20] e faz parte de um conjunto de protocolos que foram acrescentados a pilha SS7 devido aos novos requisitos de telefonia móvel, mais especificamente para atendimento as necessidades do GSM. O MAP é usado para permitir a comunicação interna entre os elementos de núcleo da rede (MSC, HLR, SGSN, etc.) para provimento de serviços, inclusive toda a parte de gerenciamento de mobilidade, *roaming* e autenticação do usuário. Além disso, o MAP também faz a comunicação inter-núcleo da rede, permitindo a comunicação entre centrais diferentes, por exemplo. Ele padroniza as mensagens que serão trocadas entre os nós da rede através de interfaces específicas. Além disso, possui uma linguagem do tipo cliente-servidor, com consultas e respostas. Na Figura 2.4, temos um exemplo de traço real com o destaque para a mensagem *"anyTimeInterrogation"* do MAP, que é usada por atacantes para descobrir a localização do usuário. Esta mensagem é interna da rede móvel e não deve ser respondida pelo elemento de rede se esta for originada em redes externas, conforme será visto na seção 2.5. Alguns dados foram omitidos por

questão de confidencialidade.

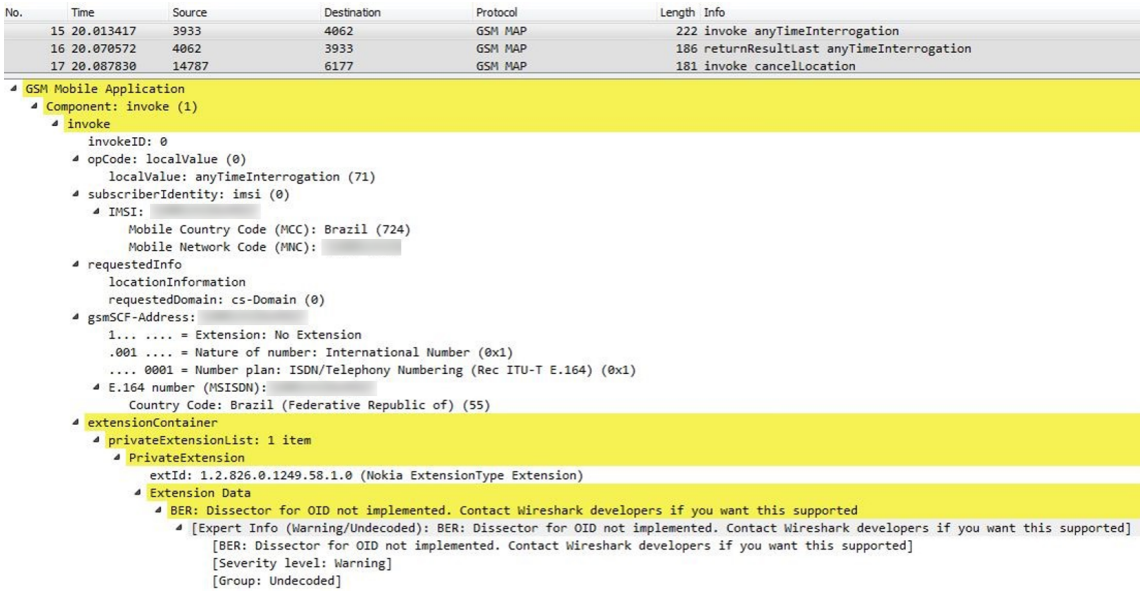


Figura 2.4: Exemplo de traço real com o conteúdo da mensagem MAP *"anyTimeInterrogation"*.

O MAP e os demais protocolos da rede móvel (tanto do núcleo quanto do acesso) são transportados via protocolo SCCP (*Signaling Connection and Control Part*). Funcionalmente, o SCCP é equivalente aos protocolos TCP (*Transmission Control Part*) e UDP (*User Datagram Protocol*) da camada de transporte do modelo TCP/IP, podendo ser orientado ou não-orientado a conexão. O SCCP utiliza o SSN (*Subsystem Number*) ao invés de portas para encaminhar os pacotes de dados para a aplicação móvel correta. Ele ainda incorpora funcionalidades de roteamento com melhorias para a camada MTP-3, fazendo uma tradução de endereços GT (*Global Title*) para PC (*Point Code*) e SSNs, além de utilizar como origem o identificador CgPA (*Calling Party Address*) e como destino, o CgPA (*Called Party Address*). Esse tipo de serviço é utilizado para atualização de localização no *roaming*. O SCCP também é utilizado para transferir consultas e respostas às bases de dados, como entre o VLR e HLR, e entre diferentes pontos de sinalização da arquitetura SS7. Na pilha SS7 (vide Figura 2.3), abaixo do MAP, há o protocolo TCAP (*Transaction Capability Application Part*), que padroniza a comunicação entre as bases de dados da rede móvel. O TCAP provê suporte a diversos serviços, tais como: portabilidade, mobilidade sem fio, serviços de rede inteligente e 0800. O protocolo CAMEL (*Customizable Applications for Mobile Enhanced Logic*) assim como o TCAP, é um protocolo de aplicação e também provê serviços para a rede móvel, como o serviço pré-pago, cujas mensagens são trocadas entre a central MSC e os elementos da rede SS7, como o SCP (base de dados). Na Figura 2.5, vemos a hierarquia entre os protocolos MTP-3, MAP, SCCP e TCAP em um traço real coletado da rede da

operadora.

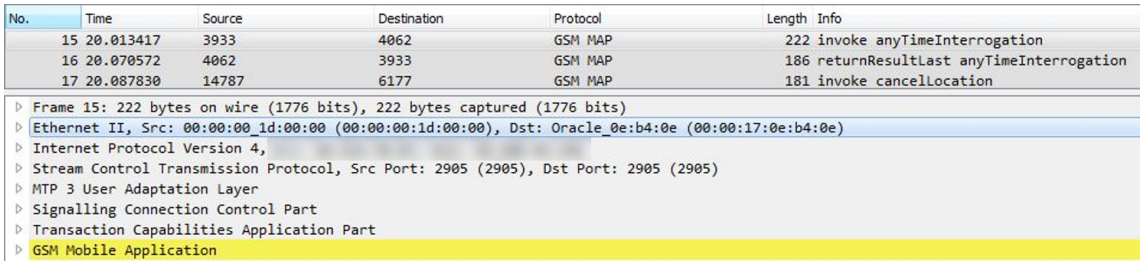


Figura 2.5: Exemplo de traço real contendo os protocolos da pilha SS7, tais como o MAP, TCAP e SCCP.

2.4 Vulnerabilidades do SS7 e Caracterização dos Ataques

O problema do SS7 surge em sua origem já que o seu projeto foi baseado em relações de interconexão confiáveis entre operadoras, que na época, eram poucas e geralmente ligadas ao governo. Esse cenário, porém, mudou nos últimos anos pois a partir do SIGTRAN e principalmente das redes IP, o SS7 tornou-se vulnerável a atacantes, deixando a rede móvel exposta. Acrescenta-se ainda o fato da explosão no número de operadoras, principalmente com o advento das operadoras móveis virtuais (*Mobile Virtual Network Operators* - MVNOs). Essa falta de preocupação com requisitos de segurança no desenvolvimento do sistema abriu um leque de possibilidades de ataques. Conforme mencionado anteriormente, a comunicação intra e inter-núcleo das redes móveis é feita através do protocolo MAP, cujas mensagens de sinalização permitem o acesso aos bancos de dados da rede (HLR/HSS e VLR) e, consequentemente, consulta às informações dos assinantes. Com isso, é possível realizar o rastreamento de chamadas de usuários, interceptação de ligações e SMS, fraudes e até indisponibilidade dos serviços de um usuário em específico ou de um elemento de rede [21].

Nas redes 4G, a situação de exposição aos ataques SS7 não é diferente. Apesar de serem redes baseadas puramente em pacotes de dados, o tráfego de voz não é suportado nativamente. Para fornecer serviços de voz integralmente sobre o LTE, ou seja, o VoLTE (*Voice Over Long Term Evolution*), é necessário introduzir na rede o IMS (*IP Multimedia Subsystem*), um tipo de *backbone* que oferece várias funcionalidades, entre elas, o suporte a voz. Porém, essa alternativa é cara, pois envolve um alto investimento em infraestrutura nas operadoras. Ademais, é uma opção a longo prazo, pois além da adequação da rede, o aparelho móvel precisa estar habilitado para originar e receber chamadas VoLTE.

Parâmetros utilizados na troca de sinalização SS7 em redes móveis			
Acrônimo	Nome	Comprimento	Descrição
GT	*Global Title*	15 dígitos (máx.)	É um número telefônico composto pelo PC+SSN. Identifica unicamente um nó na rede móvel.
MSISDN	*Mobile Station International Subscriber Directory Number*	15 dígitos (máx.)	Número do telefone do usuário. Composto pelo código do país + código de área + nº do assinante.
MCC	*Mobile Country Code*	3 dígitos	Compõe o IMSI identificando o país.
MNC	*Mobile Network Code*	2 a 3 dígitos	Compõe o IMSI identificando a operadora.
NDC	*National Destination Code*	3 dígitos	Compõe o MSISDN identificando a localidade.
MSIN	*Mobile Station Identification Number*	10 dígitos (máx.)	Compõe o IMSI atribuindo um número único ao assinante.
IMSI	*International Mobile Subscriber Identity*	15 dígitos (máx.)	Formado pelo MCC+MNC+MSIN. Identifica o assinante unicamente na rede e é gravado no SIMCard.
IMEI	*International Mobile Equipment Identity*	15 dígitos (máx.)	Identifica uma estação móvel de forma única. É armazenado permanentemente no aparelho durante sua fabricação. Formado pelo número de série, fabricante, país e tipo de instalação.
PC	*Point Code (Originating ou Destination)*	24 bits	Similar a um endereço IP. Identifica um Ponto de Sinalização na camada MTP-3 da pilha SS7.
SSN	*SubSystem Number*	3 decimais ou 8 bits	Similar às portas TCP/UDP. Identifica os nós da rede para o protocolo SCCP. Ex.: HLR (6), VLR (7), MSC (8), etc.
TSMI	*Temporary Subscriber Mobile Identity*	4 octetos	Pseudônimo para mascarar a identidade do assinante. Usado pelo VLR a fim de que o IMSI não seja identificado. Utilizado no procedimento de atualização de localização.

Tabela 2.1: Identificadores de rede utilizados na sinalização SS7 para troca de mensagens nas redes móveis.

Ainda que a chamada seja inteiramente em LTE via IMS, se o assinante sair da área de cobertura 4G, a chamada será interrompida. Com isso, para que haja uma garantia de continuidade da ligação onde a cobertura do LTE não alcança, torna-se necessário realizar um *fallback* (CSFB) para redes anteriores, 3G ou 2G. Isso mostra que mesmo as redes mais novas não estão isentas de sofrer com as consequências do relaxamento da segurança do sistema SS7. Outrossim, o LTE utiliza o protocolo Diameter ao invés do SS7, que apesar de prover maior segurança de rede, ainda não foi implantado totalmente pois muitos nós são híbridos, com suporte a ambos protocolos. Todos estes argumentos corroboram o atual problema de segurança trazido pelo SS7, ainda sem perspectiva de solução, ou pelo menos enquanto as futuras redes 5G não forneçam totalmente os serviços oferecidos pelas atuais gerações de telefonia móvel.

A fim de obter uma melhor compreensão dos ataques que serão explicitados na próxima seção, a Tabela 2.1 apresenta os principais endereços de rede que são obtidos e utilizados para execução dos ataques.

2.4.1 Capacidades do Atacante

Há diversas possibilidades para execução de um ataque em SS7. O atacante também pode combinar as diferentes formas a fim de atingir melhores resultados. Uma delas é através da interface aérea, que é o meio de transmissão de dados entre a estação móvel e a BTS. Diferentes canais lógicos são alocados para esta comu-

nicação, comuns e dedicados. Apenas os canais dedicados possuem autenticação e um atacante com os recursos necessários pode se beneficiar desta fragilidade.

Dentre os canais comuns, pode-se destacar o BCCH (*Broadcast Common Control Channel*) e o PCH (*Paging Channel*). Toda vez que um aparelho celular é ligado, ele se encontra no modo *"Idle"* e aguarda de forma passiva as informações da rede para acessá-la. O BCCH é usado pela BTS para enviar em *broadcast* estas informações da rede, como a identificação da célula (*cell ID*); a área de localização da célula; o MCC, o MNC e as frequências usadas pelas células vizinhas a fim de otimizar a alocação dos canais para os terminais móveis. Já o PCH é utilizado para contatar as estações móveis em modo "*Idle*"na recepção de uma mensagem ou uma ligação de voz. Para isto, o equipamento do usuário continuamente envia atualizações de localização para alertar a sua disponibilidade na rede. Em contrapartida, a rede envia uma mensagem de *Paging* em *broadcast* para as células localizadas na última área reportada pelo assinante, a fim de encontrá-lo para entregar os serviços à ele. Para esta consulta, a BSC procura pelo TSMI do assinante, um identificador temporário que é dinamicamente atribuído cada vez que o móvel se registra na rede e um canal dedicado é alocado. Isso evita que o IMSI seja divulgado através do canal PCH, omitindo a verdadeira identidade do usuário.

De posse de um equipamento de rádio passivo ou um rádio definido por *software* (SDR), o atacante pode capturar todas as informações divulgadas pelos canais comuns (BCCH e PCH), já que as informações trafegadas por meio destes ocorre sem criptografia. O BladeRF x40[1], mostrado na Figura 2.6 é um exemplo dos rádios SDR disponíveis no mercado a preço acessível que pode ser utilizado como farejador passivo para escuta na interface aérea. Karsten Nohl, da Security Research Labs, utiliza este dispositivo, combinado com um programa gerador de consultas em SS7 baseado em Linux para demonstrar a execução de um ataque de interceptação de SMS [22]. Outra opção semelhante é o uso de USRP (*Universal Software Radio Peripheral*) em combinação com um *software* livre, como o GnuRadio[2], que oferece um ambiente de desenvolvimento de aplicações, podendo também servir como um ouvinte da comunicação da rede de acesso.

A partir da alocação de um canal dedicado para o usuário, a comunicação passa a ser criptografada para que o usuário possa enviar seus dados à rede, como o IMSI, de forma segura. Rupprecht *et al.* exibe uma tabela com os algoritmos usados nesta comunicação, por exemplo, para o GSM tem-se o A5/1, A5/2, A5/3 e GEA4, dentre outros, a maioria de 64 bits de tamanho da chave [23]. Contudo, esta criptografia no GSM mostrou-se frágil, já que foi possível quebrar a segurança da comunicação que baseava-se nos algoritmos A5/1 e A5/2, por exemplo [24–26].

[1]https://www.nuand.com/product/bladerf-x40/
[2]https://www.gnuradio.org/about/

Figura 2.6: Exemplo de Rádio Definido por *Software* chamado bladeRF x40.

No entanto, ainda que novos algoritmos mais robustos sejam desenvolvidos para o GSM, há ainda a questão da ausência de autenticação mútua. Isto é, existe apenas a autenticação unilateral da estação móvel para a rede e presume-se que a outra parte (estação base) que está solicitando dados para o usuário seja legítima. Um atacante que possua um rádio ativo (chamados de *"Cell-Site simulators"* ou "*IMSI Catcher"*) como o StingRay I/II[3], Engage GI2[4] ou outro DRT (*Digital Receiver Technology*)[5] pode executar um ataque MiM ("*Man-In-The-Middle*"), estabelecendo uma conexão intermediária entre o usuário e a rede. Também pode ser utilizado um *software* de código aberto, como o OsmocomBB[6], que é instalado em um celular e conectado a um PC para agir passiva e ativamente na rede GSM. Para tal, o atacante com algum destes recursos realiza continuamente o *broadcast* do MCC e MNC da rede, personificando uma estação base e ocultando as frequências das BTS's vizinhas para que o assinante não se conecte a uma estação base real. O aparelho celular do usuário por sua vez, estando em modo "*Idle*", se incorpora à BTS mais próxima, com maior intensidade de sinal. Do ponto de vista da rede, o atacante se camufla personificando o usuário. Uma vez conectado à estação falsa, as mensagens trocadas entre a estação base verdadeira e o assinante é feita por meio do atacante, que age de forma transparente, repassando as mensagens de um lado a outro e consequentemente, obtendo informações do assinante, como IMSI, IMEI, localização etc. O atacante também pode desviar o tráfego para um modem específico. A partir do 3G, a autenticação mútua foi implantada utilizando algoritmos robustos de 128 bits de tamanho da chave [27], porém os atacantes realizam o rebaixamento de tecnologia para o 2G.

[3]https://theintercept.com/surveillance-catalogue/stingray-iii/

[4]https://www.documentcloud.org/documents/885760-1278-verint-product-list-engage-gi2-engage-pi2.html

[5]https://www.documentcloud.org/documents/2185450-digital-receiver-technology-presentation.html

[6]https://bb.osmocom.org/trac/

Conforme visto anteriormente, o SS7 não foi projetado para atender aos requisitos de segurança exigidos atualmente. Portanto, qualquer mensagem MAP ou de outro protocolo da pilha SS7 destinada à rede de núcleo móvel ou a algum ponto de sinalização SS7 é respondida, assumindo-se que o emissor é um nó de rede legítimo. Obter acesso à rede SS7 e começar a gerar mensagens não é tão trivial mas é viável. Por exemplo, é possível alugar ou adquirir diretamente um *hub* SS7 que possua acesso à rede da operadora como as soluções da Comfone[7] e da FGT[8]. Essa utilidade é oferecida geralmente para que MVNOs ou pequenas operadoras tenham acesso à rede SS7 de uma grande provedora de serviços de telecomunicações, a fim de permitir a terceirização e extensão de serviços, como *Roaming* e SMS. Uma vez de posse do *hub*, além do acesso à rede, é obtido um GT (*Global Title*) e um acordo de *roaming*, se este for o serviço a ser oferecido. Um atacante pode comprar este acesso e passar a gerar mensagens em SS7 legítimas à rede. Outra forma de obtenção de entrada à rede SS7 é através da compra de femtocélulas (ou HNB (*Home NodeB*) no 3G), que são um ponto de acesso do usuário à rede da operadora e já demonstraram vulnerabilidades à ataques, permitindo aos atacantes a obtenção do IMSI, IMEI e até mesmo a interceptação de chamadas e SMS [28]. Ademais, também está disponível uma ferramenta de rastreamento de localização de usuários, chamada "SkyLock"desenvolvida pela Verizon[9], que foi veiculado no jornal *The Washington Post*[10]. A única informação a ser inserida no programa para consulta sobre a localização de um assinante é o seu número de telefone (MSISDN).

Todas estas alternativas para execução de ataques em redes móveis podem ser feitas de forma combinada a fim de potencializar o impacto do ataque.

2.4.2 Principais Tipos de Ataques

Avizienis *et al.* definem os pilares para a segurança de rede [29], dentre as quais destaca-se: disponibilidade, confiabilidade, confidencialidade, integridade e manutenibilidade. Cada ataque em SS7 tem como objetivo final a infração de alguma destas premissas. Pode-se dividir os ataques em sinalização SS7 nas seguintes categorias: Rastreamento, Interceptação, Fraude e Negação de Serviço.

Rastreamento: Consiste em seguir o posicionamento da estação móvel em tempo real, permitindo localizar a vítima a qualquer momento. Tem como objetivo a violação da confidencialidade da rede, já que a localização do usuário é uma informação de privacidade, não devendo ser autorizada a sua divulgação para terceiros. É um ataque de difícil mitigação, já que o assinante precisa enviar mensagens de atua-

[7] https://www.comfone.com/services/sponsored-roaming-solution
[8] http://fgtglobal.com/
[9] http://apps.washingtonpost.com/g/page/business/skylock-product-description-2013/1276/
[10] https://wapo.st/2lSe5nQ

lização de localização periódicas para a rede a fim de demonstrar sua disponibilidade e localidade, o que é explorado por atacantes devido a insegurança nos canais comuns usados para esta comunicação.

Interceptação: A intenção do atacante é ler dados originalmente enviados para outro usuário. A partir das informações obtidas através do ataque de rastreamento, é possível iniciar o ataque de interceptação. De posse destas informações, o atacante pode gravar conversas, ler senhas e obter informações sobre as atividades dos usuários, apenas reencaminhando as mensagens SMS ou chamadas para si próprio, que supostamente é um elemento legítimo da rede móvel. Assim como o ataque de rastreamento, o objetivo final da interceptação é a quebra de sigilo da privacidade do usuário. Apesar das informações serem trafegadas por um canal dedicado, no GSM há a possibilidade de execução do ataque MiM, por exemplo, que usa vulnerabilidades da falta de autenticação mútua e de segurança do SS7. Nas tecnologias mais recentes, o atacante pode migrar para o 2G e executar a interceptação.

Fraude: O atacante tem como objetivo obter vantagens financeiras de um usuário da rede. Um exemplo é a alteração das configurações de encaminhamento de chamadas usando o protocolo MAP para que o custo de chamadas seja arcado por outro usuário. A motivação desse ataque é utilizar serviços mais sofisticados tais como números de telefones usados para jogos de azar, *chats*, número VIP, etc., sem ser tarifado por isso.

Negação de Serviço: Consiste em interromper o serviço dos assinantes (por exemplo, a realização de chamadas e de SMS) ou tornar algum elemento de rede indisponível. Esse ataque visa comprometer a disponibilidade da rede, a integridade e a manutenibilidade.

A seguir, os ataques serão descritos com mais detalhes.

O ataque de rastreamento é o mais simples de se executar, bastando que o atacante tenha acesso a alguma ferramenta de rastreamento em tempo real, como o SkyLock. Outra possibilidade é através do ataque MiM e envio de mensagens MAP. O atacante com este tipo de acesso e número do assinante, se personifica como um elemento legítimo de rede e envia uma mensagem *"anyTimeInterrogation"* (ATI) do protocolo MAP para a base de dados (HLR) da rede de origem do assinante, solicitando a localização do usuário (*cell ID*) e o IMEI. O HLR contacta a central da rede visitada questionando sobre o ID da célula. Isto é feito por meio da mensagem *provideSubscriberInfo* (PSI) para o VLR em que o assinante se encontra atualmente. Por fim, o VLR realiza o processo de *Paging* para obter as informações do usuário a partir da estação base e as reencaminha para o HLR, retornando a informação de localização para o atacante. Este processo é exemplificado na Figura 2.7.

Para os ataques de interceptação, é necessário que o atacante obtenha previamente alguns dados da rede, como um identificador válido (GT) de um elemento,

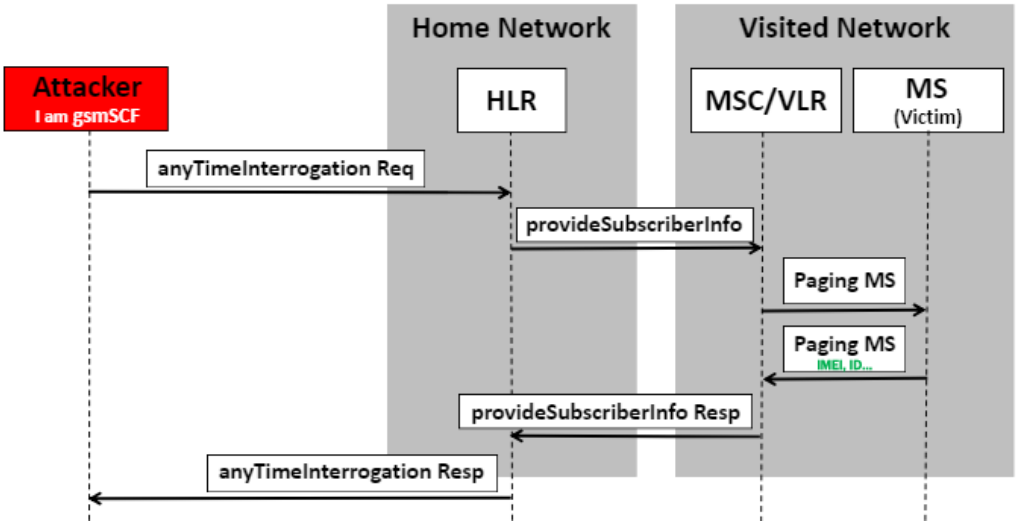


Figura 2.7: Troca de mensagens MAP no ataque de rastreamento.

por exemplo, do SMSC para interceptação de mensagens SMS. Isto pode ser feito através de escuta da interface aérea, conforme abordado anteriormente. De posse desta informação e conhecendo o número do assinante, o atacante habilitado a gerar mensagens MAP requisita as informações de roteamento do assinante, como por exemplo, o IMSI. Como os elementos de rede respondem a qualquer mensagem MAP válida, o banco de dados (HLR) atende a solicitação com os dados do assinante. Esta troca de mensagens pode ser vista na Figura 2.8. Em azul, dentro da mensagem MAP enviada pelo atacante, tem-se o que foi obtido na interface aérea (GT); em verde, a informação já conhecida pelo atacante e em vermelho, a informação que o mesmo deseja obter.

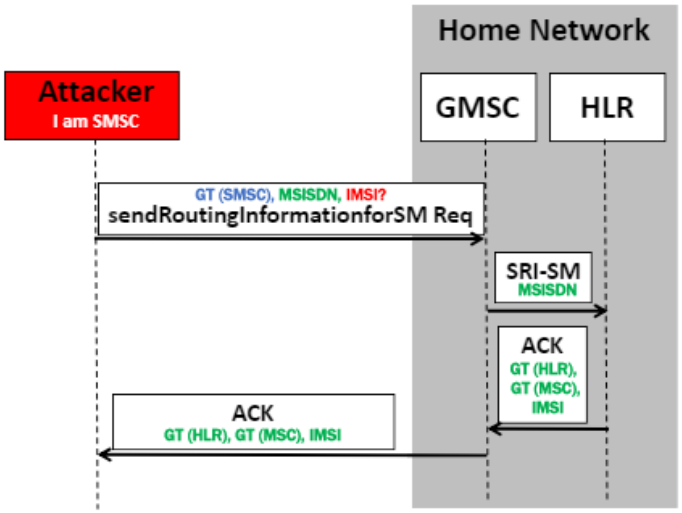


Figura 2.8: Troca de mensagens MAP no ataque de interceptação de SMS.

Após conseguir o IMSI do usuário, o atacante precisa de um novo GT válido para personificar uma central MSC. Conforme abordado na seção anterior, este

identificador válido (GT) pode ser adquirido por meio da operadora, através da compra de um *hub* SS7. O atacante, por sua vez, atua agora como central MSC e envia uma mensagem MAP diretamente para o HLR, alterando a localização atual do assinante para a nova MSC, exemplificado na Figura 2.9.

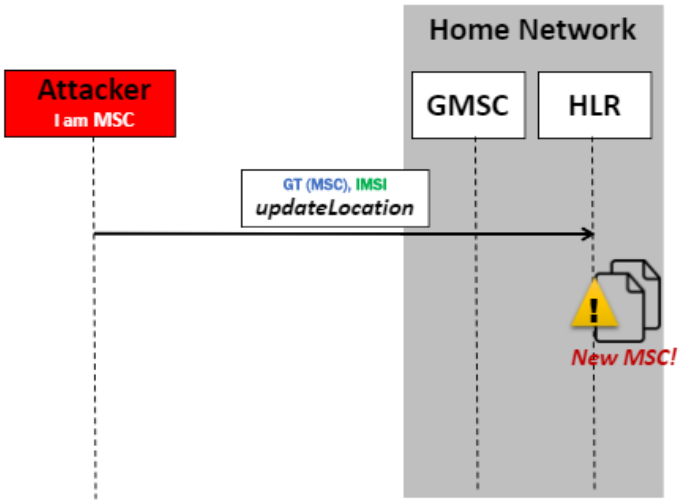


Figura 2.9: Troca de mensagens MAP no ataque de interceptação de SMS, com o atacante como MSC enviando a mensagem *updateLocation*, que atualiza a localização atual do assinante.

Quando um SMS chega na rede, o SMSC procura o assinante no banco de dados a fim de entregar a mensagem. O HLR responde que a localização atual do assinante em questão é a MSC do atacante, o qual declarou possuir o usuário em sua rede. Portanto, todos os SMS's enviados para este assinante serão redirecionados para o atacante, vide Figura 2.10.

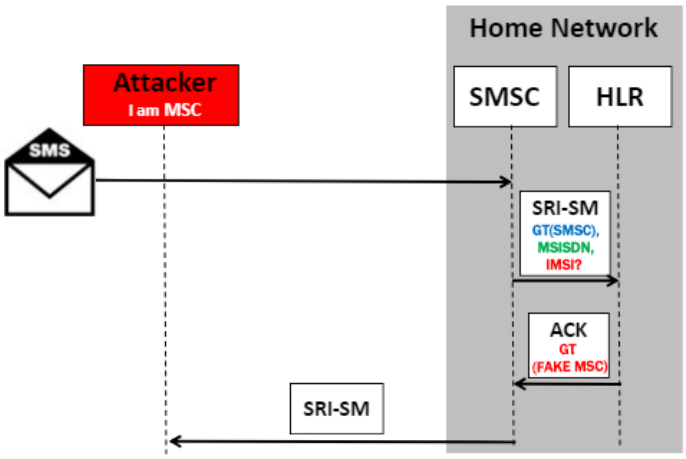


Figura 2.10: Atacante obtém o desvio do SMS para si e pode armazenar, alterar e/ou encaminhar posteriormente o SMS ao assinante.

Para o ataque de negação de serviço, além do acesso à rede SS7 e do número de telefone do usuário alvo, o atacante também precisa obter previamente o IMSI

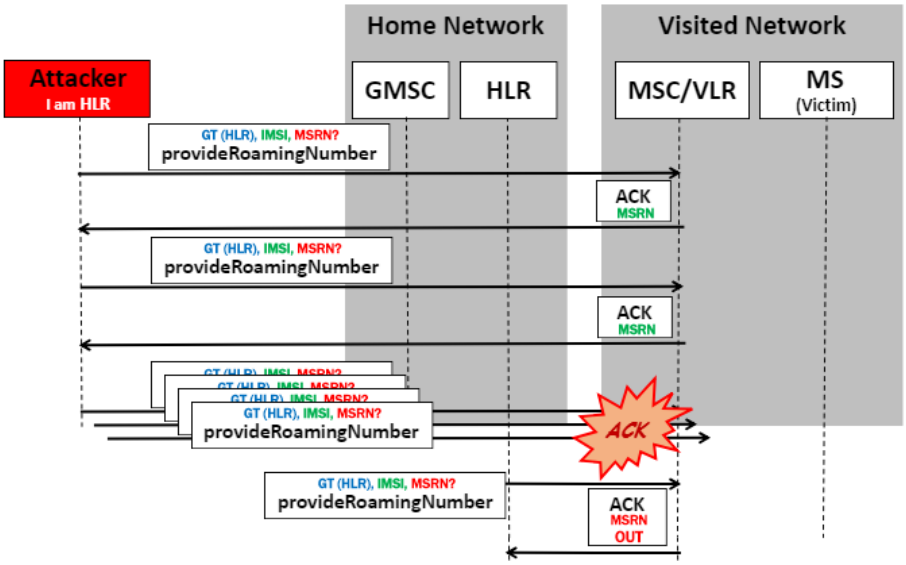


Figura 2.11: Troca de mensagens MAP no ataque de negação de serviço.

do assinante e o *Global Title* (GT) da central MSC e do banco de dados VLR. A identificação GT pode ser conseguida facilmente na Internet, através do documento "IR.21", que padroniza os acordos de *roaming* entre as operadoras. Fazendo uma busca simples, foi possível obter todos os GTs de uma grande operadora que atende o Brasil, além de outras informações sensíveis, como os fornecedores de cada elemento de rede e a localização de cada plataforma. Após obtenção das informações supramencionadas, o atacante envia a mensagem MAP *"deleteSubscriberData"* para a MSC/VLR, que então remove todos os serviços habilitados para o assinante em questão. O atacante pode conseguir o mesmo feito enviando a mensagem *"cancelLocationreq"*, que remove a conexão do assinante com a rede, impossibilitando chamadas e SMS's de serem entregues a este usuário.

Para que um elemento de rede fique indisponível, o atacante com acesso à rede SS7 simula várias requisições de *Roaming*, conforme ilustrado na Figura 2.11. A cada requisição, um número de *roaming* é provido para o falso elemento de rede. Para isto, o atacante envia uma mensagem MAP *"provideRoamingNumber"* com seu GT, pedindo o número MSRN (*Mobile Subscriber Roaming Number*) associado ao IMSI da vítima. O MSC responde informando este número. O atacante então repete a requisição por diversas vezes, inundando a central de comutação, até o momento em que não haja mais *Roaming Number* disponível. Quando um banco de dados (HLR) real fizer esta mesma requisição, a MSC retornará que esgotaram os números de *roaming* disponíveis, provocando a indisponibilidade da estação base pertencente à MSC ao receber ligações.

2.5 Principais Contramedidas

A IR.82 é um documento publicado pelo GSMA em 2016 que trata de implementações de segurança na rede SS7, com sugestões para o tratamento da vulnerabilidade do SS7 [30], que é a confiança mútua entre os elementos e atendimento às solicitações sem checagem da procedência do emissor. Para isto, são definidas 3 categorias de mensagens MAP e CAP (*CAMEL Application Part*), para as quais recomenda-se a filtragem no envio e na recepção dentro da rede em nós terminais (nós do núcleo da rede móvel tais como MSC, HLR, VLR, etc.) e nós de trânsito (pontos de sinalização da rede SS7, por exemplo, STP). Para cada categoria, também é proposto um método de filtragem. São elas:

- Categoria 1: Composta pelas mensagens MAP que devem ser recebidas dentro da mesma rede, ou seja, mensagens internas. Apenas deve-se aceitar uma mensagem deste tipo advinda de redes externas se houver um contrato que estabeleça este recebimento. A mensagem ATI ou *"AnyTimeInterrogation"*, é uma delas, que deveria ser recebida e tratada apenas se o originador for o HLR da rede de origem. Para este tipo de categoria, a recomendação é que deve ser adotada uma política de filtragem na borda da rede da operadora através de *firewalls*.

- Categoria 2: São mensagens MAP de *roaming* provenientes da rede de origem do assinante para a rede visitada. Por exemplo, um assinante local do RJ está viajando em SP e o banco de dados de origem HLR do RJ faz uma consulta à rede de SP a fim de checar a localização do assinante. Para prevenir ataques neste tipo de categoria, a orientação do GSMA é checar a fonte do pacote da mensagem, verificando se o IMSI do assinante visitante pertence ao HLR da rede de origem.

- Categoria 3: Mensagens MAP de *roaming* a partir da rede visitada do assinante para a rede de origem. É parecido com o caso anterior, porém, a consulta parte do VLR de SP para o HLR do RJ, por exemplo. É mais difícil de aplicar filtros, pois a mensagem pode ser originada de qualquer lugar, um atacante pode afirmar que o assinante está em sua rede. A rede de origem deve comparar a última informação de localização do usuário com a fonte da mensagem MAP recebida e caso as informações não coincidam, deve-se avaliar se foi possível neste espaço de tempo que o assinante tenha viajado para a nova localidade informada. Se for concluído que a localização atual não é razoável, deve-se abortar a mensagem.

A solução definitiva para os ataques em SS7 seria a completa substituição do sistema. No entanto, esta solução não é factível para os próximos anos. Portanto,

algumas alternativas vêm sendo estudadas, como a filtragem de mensagens MAP, já abordada acima. Além deste protocolo, também há iniciativas para bloqueio de mensagens CAP, SCCP e SMS [30].

O 3GPP também atuou de forma a minimizar os ataques SS7 de interceptação de SMS, publicando em 2007 uma modificação para o tratamento de SMS na rede. Anteriormente, o SMSC enviava uma consulta ao HLR para descobrir a central cujo assinante destinatário estava alocado. Uma vez descoberta, é feito o encaminhamento do SMS para esta central e desta para o assinante. Portanto, um atacante personificando a função de SMSC, poderia utilizar este meio para interceptar mensagens SMS. A partir das informações coletadas (como endereço GT da central e IMSI do assinante), o atacante pode então se passar pela central destinatária assim que um SMS é recebido, não entregando a mensagem ao destino ou agindo de forma transparente para obter a leitura do SMS. A partir do "*SMS Home Routing*", um roteador SMS é colocado entre a rede de origem e a rede de destino, e, portanto, o encaminhamento de mensagens não é mais realizado de forma direta, dependendo de um intermediário (roteador) e todas as trocas de mensagens limitam-se à rede de origem. Além disso, o IMSI não é trocado entre o roteador e a rede de destino, utilizando um identificador de correlação a fim de ocultar os dados sensíveis da rede. A grande questão é que nem todas as operadoras implantaram o *"SMS Home Routing".*

Para o caso de personificação de elementos de rede, existem iniciativas no sentido de detecção da presença de anomalias na rede, dentre aplicativos e sensores [23]. Por exemplo, Karsten Nohl apresentou um aplicativo chamado "SnoopSnitch"desenvolvido pela Security Research Labs para a plataforma Android [22]. Além de detectar a presença de "*IMSI Catchers*", ele mostra se a rede da operadora que o usuário possui está sofrendo algum tipo de ataque em SS7, como o de interceptação em SMS. O aplicativo permite que seja configurado um alarme para avisar que um ataque em SS7 foi detectado. O SnoopSnitch funciona apenas com chipsets da Qualcomm e o usuário precisa ter o celular habilitado o modo *root* para permitir o escaneamento da rede da operadora. A empresa não deixa claro a forma de detecção na rede, mas supõe-se que o aplicativo torna o celular um rádio passivo e ativo, executando ataques em SMS para verificar a vulnerabilidade da operadora e verificando passivamente as transmissões do canal de rádio e potência do sinal das estações-base próximas para detectar a presença de *Cell Site Simulators* na rede. Os dados coletados pelo aplicativo podem ser enviados para a base de dados que compõe o GSMMap, um mapa *online* que mostra a exposição das redes das operadoras a ataques em SS7 em nível global.

Outra proposta para tratar o caso dos rastreadores de IMSI é a melhoria da autenticação da rede utilizando um pseudônimo para este identificador, chamado de

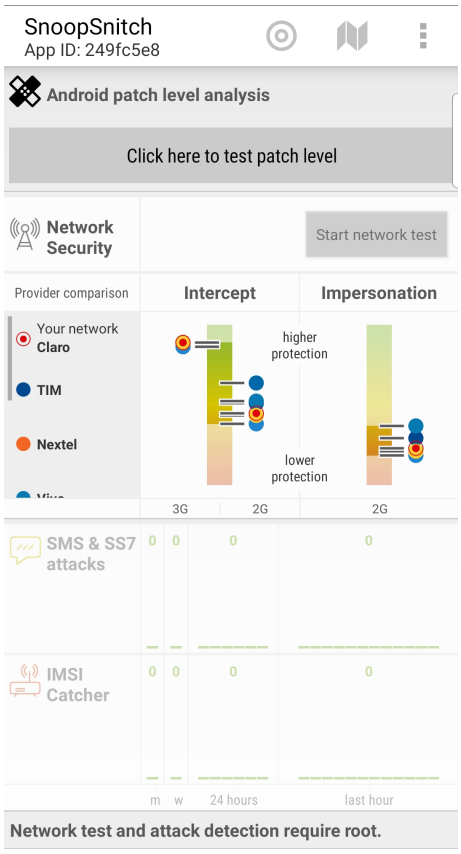


Figura 2.12: Tela do aplicativo *SnoopSnitch*, que detecta ataques em SS7 na rede do usuário.

PMSI (*Pseudo Mobile Subscriber Identity*) [31]. O PSMI é gerado aleatoriamente e tem como objetivo esconder a verdadeira identidade do assinante para obter confidencialidade no canal de comunicação.

Capítulo 3

Caracterização dos Dados Reais

Uma grande operadora de telecomunicações que opera no Brasil verificou a vulnerabilidade de sua rede através de testes. O objetivo dos testes foi levantar as possíveis fraudes que poderiam estar ocorrendo através das redes de sinalização SS7. Os detalhes da implementação encontram-se na próxima seção e, posteriormente, serão apresentados os resultados encontrados a partir dos dados coletados.

3.1 Coleta dos Dados

A Figura 3.1 mostra o cenário de testes utilizado. A sinalização SS7 recebida de operadoras internacionais é enviada para o destino interno da rede. Os pontos de transferência de sinalização da rede (STPs) duplicam o tráfego recebido de redes externas e enviam uma cópia para o "STPCRT". Este tráfego bruto de *roaming* internacional é então espelhado para um servidor Linux disponibilizado por um fornecedor com função de *firewall* de sinalização, disponível no mercado, que faz a detecção de ameaças em SS7. Neste *firewall* foram aplicados filtros específicos para identificação das ameaças destinadas à operadora e suas parceiras, como por exemplo, a configuração do MCC para "724"(Brasil) e do MNC e NDC, para o valor correspondente ao da operadora e localidade, para que toda a sinalização enviada para IMSI's iniciados pelo prefixo "7240X"(onde X é o número da operadora) seja um ponto de observação. Para tal, foi utilizado o sistema de detecção de mensagens suspeitas de fraude (*IDS – Intrusion Detection System*) que trabalha em conjunto com o *firewall* de sinalização. Neste sistema foram criados perfis de classificação para definição da ocorrência de um ataque. Esta definição pode ser subjetiva, pois nem todos as classificações correspondem a um ataque de fato, gerando um falso positivo. Por exemplo, o fornecedor do *firewall* pode estabelecer que uma mensagem com sintaxe errada seja um ataque ou que uma mensagem de Categoria 1 deve ser bloqueada. Porém, esta pode estar declarada como exceção no contrato de *roaming* da operadora. Ou seja, as classificações devem ser definidas em conjunto com a

operadora, a análise dos ataques é feita de forma específica para cada caso. Portanto, neste capítulo, a palavra "ataque"será substituída por "ameaça", já que nem todos os ataques lograram sucesso. Com isso, é preciso conhecer as peculiaridades da operadora a fim de não classificar um ataque de forma equivocada. Por exemplo, a provedora de serviços em questão utiliza o protocolo SMPP (*Short Message Peer-to-Peer*) e um SMS *Gateway* para encaminhamento de mensagens SMS, portanto, na análise dos dados, não foi encontrada uma ameaça de interceptação de SMS.

Previamente, o tráfego bruto foi capturado para análise pelo programa Wireshark, porém esta opção não se mostrou viável pelo alto fluxo de mensagens, que continuamente saturava o *buffer* de armazenamento das mesmas. Portanto, conforme citado, a solução adotada foi utilizar um *firewall* do mercado, que além de possuir maior capacidade de armazenamento, é uma ferramenta de detecção e mitigação de ameaças em SS7. Basicamente, ele utiliza uma ferramenta *Analytics* para análise dos dados para encontrar anomalias. Alejandro Corletti, da DarFe, mostrou que é possível fazer a análise de dados em SS7 apenas com o Wireshark e o uso de um *software* livre, como o Snort[1] [32].

Após a implementação do servidor pelo fornecedor, os dados foram disponibilizados para acesso pela Internet, por SSH (*Secure Shell*) ou por interface gráfica em um navegador *Web*. A partir do acesso a esta interface, foi possível extrair planilhas diárias, semanais e mensais com as informações sobre as ameaças. O acesso aos dados foi autorizado pela operadora bem como os dados divulgados nesta dissertação. A partir da coleta, foram desenvolvidos códigos no MATLAB para caracterização dos dados.

A coleta aconteceu durante o período de 5 meses, entre 15/12/2018 e 15/05/2019, totalizando 150 dias de observação. As subseções que seguem expõem e analisam o tráfego coletado.

3.2 Resultados obtidos

Este capítulo tem como objetivo caracterizar o problema de ameaças aos protocolos da pilha SS7 enfrentado pelas operadoras de telefonia móvel. Para isso, é feita a análise de um traço de 150 dias contendo dados de ameaças a uma rede de um provedor de serviço brasileiro de grande porte. Os resultados mostram os principais tipos de ameaças observadas, a quantidade de ameaças sofridas pela operadora neste período bem como a intensidade das mesmas, a distribuição da origem e destino das ameaças e a frequência de transições entre ameaças consecutivas.

[1] https://www.snort.org/

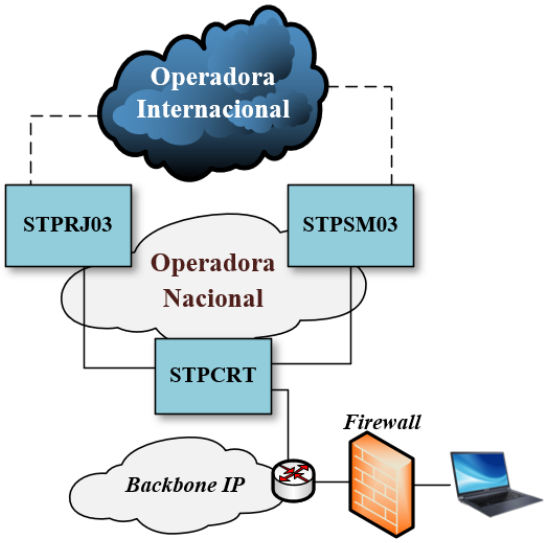


Figura 3.1: Topologia utilizada nos testes na rede da operadora. O tráfego recebido das operadoras internacionais é replicado e copiado para um *firewall*.

3.2.1 Intensidade das ameaças

A Figura 3.2 mostra a quantidade total de ameaças identificadas durante o período de avaliação. Essa quantidade aproxima-se da média na maior parte do tempo, que é de 21.900 ameaças/dia. Aconteceram picos nos meses de Dezembro e Março, mais especificamente nos feriados de Natal e Carnaval, provavelmente pela chance de sucesso do atacante ser maior em períodos fora do expediente da operadora ou em períodos de maior uso por parte dos usuários. Ainda, um pico aconteceu entre os dias 17 e 19/03, onde verifica-se que no dia 18/03, as ameaças chegaram a ser disparadas 118.162 vezes. A escolha de tais datas carecem de maiores investigações de pesquisa. Não foi possível identificar a real quantidade de ameaças bem sucedidas, pois esta informação é confidencial e, portanto, os dados da Figura 3.2 representam as tentativas de ameaças realizadas contra a rede da operadora.

As Figuras 3.3 e 3.4, mostram as classes de ameaças encontradas nos dados analisados. Estas figuras complementam-se na apresentação dos tipos de ameaças mencionados na Subseção 2.4.2. A categoria "Interceptação" é a principal, apresentando 40,9% do total de ameaças. Nessa categoria, a rede da operadora é utilizada como meio para obtenção de dados dos clientes, comprometendo a privacidade dos mesmos. Em sequência, com 29,99% e 28,82%, encontram-se as categorias "Negação de Serviço" e "Fraude". Com 0,26% e 0,03%, tem-se respectivamente as classes "Negação de Serviço da Rede" (na qual os serviços de rede ficam indisponíveis) e "Rastreamento".

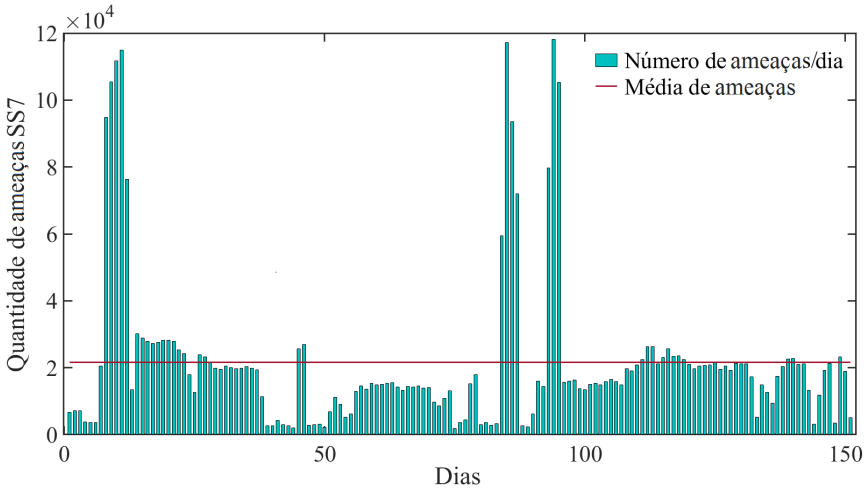


Figura 3.2: Quantidade de ameaças no período de 15/12/18 à 15/05/19.

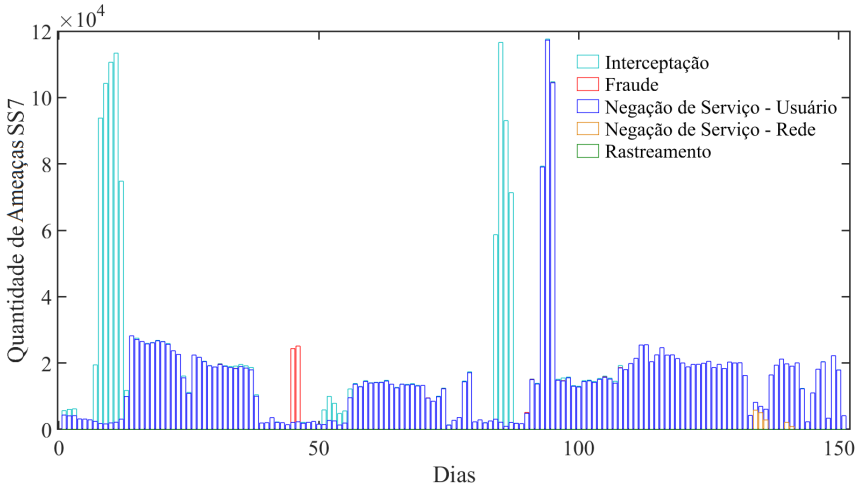


Figura 3.3: Distribuição diária das ameaças encontradas na rede da operadora.

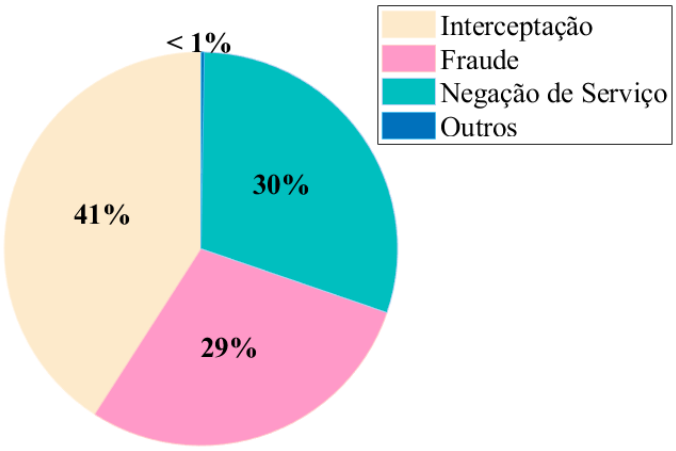


Figura 3.4: Tipos de ameaças encontradas nos dados extraídos da rede da operadora.

3.2.2 Ameaças mais frequentes

A partir dos dados analisados, é possível discriminar quais ameaças dentro das categorias já mencionadas são as mais frequentes na rede. As seis principais ameaças estão descritos na Figura 3.5. São eles: *Fast Relocation, Abnormal TCAP Handshake Aborted, Fast Relocation with SMS, SendRoutingInfoForSM Abnormal, CL Completed* e *PSI Completed*.

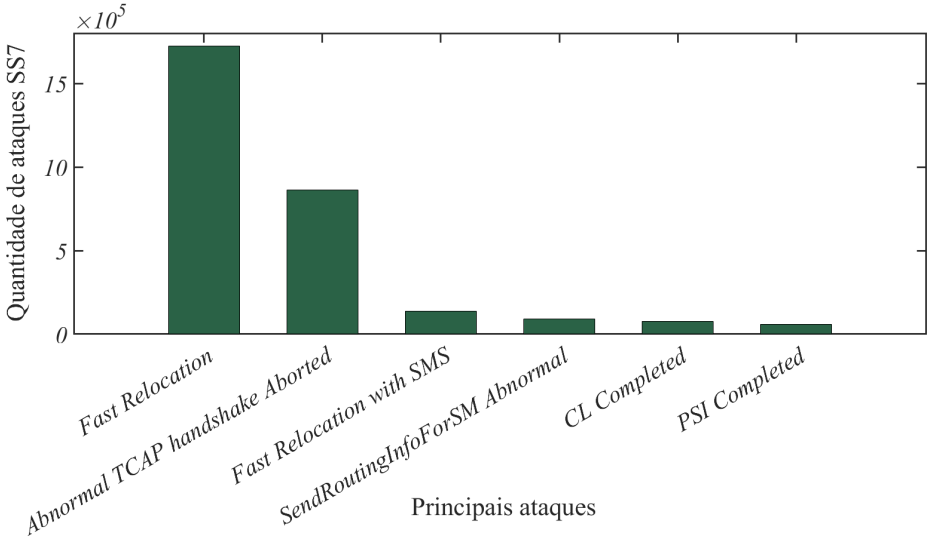


Figura 3.5: Ranqueamento das seis principais ameaças.

Uma ameaça pode ser considerada como *Fast Relocation* quando o tempo de transição (ou realocação) de um assinante está abaixo de um valor aceitável. Dessa forma, se um assinante se move de um país para o outro utilizando o serviço de *roaming* internacional, para manter o seu número habilitado para receber chamadas, enviar SMS e consumir o pacote de dados, a operadora deve possuir a informação de localização e estampa de tempo da última vez em que o assinante esteve conectado à rede e também a localização e informação de data e hora atuais. Se a diferença entre as estampas de tempo não estiver coerente (levando em consideração a distância entre os países e fuso horário), esta movimentação na rede é considerada uma ameaça, pois o atacante pode ter transferido o assinante para a sua rede ilegítima. Já a ameaça *Fast Relocation with SMS* se dá quando a operadora recebe uma mensagem SMS de um assinante que se encontra em outro país na situação de *roaming*. Essas ameaças podem ser consideradas como "Negação de Serviço" (onde os serviços se tornarão indisponíveis para o assinante, caso haja a exclusão do usuário na rede da operadora), "Fraude" ou "Interceptação".

A ameaça *Abnormal TCAP handshake Aborted* é classificada dentro da categoria "Interceptação" e ocorre quando uma mensagem TCAP (*Transaction Capability Application Part*) é recebida com ausência de alguns parâmetros. Essa mensagem

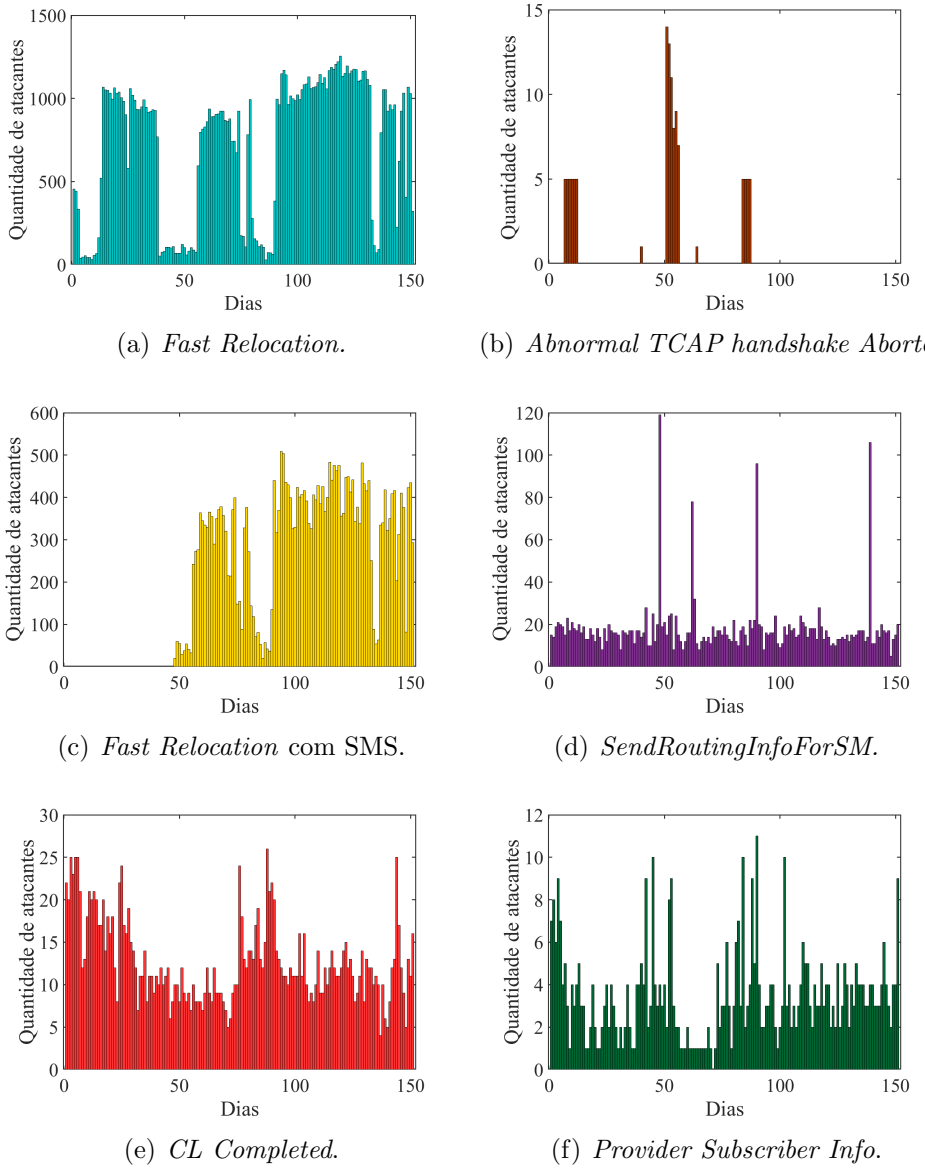
portanto é ilegítima e considerada pela rede como uma ameaça. Por padrão, ao receber essa mensagem, a mesma é descartada por não obedecer os critérios do 3GPP. A ameaça *Send Routing Information for Short Message Abnormal* é uma mensagem MAP de requisição para recuperar dados do usuário, como o IMSI e localização. Esta solicitação é considerada uma ameaça quando não é precedida de uma mensagem de envio de SMS, para o qual se destinaria. A mesma é abortada pela rede e categorizada como "Fraude", "Interceptação" ou "Negação de Serviço" para o usuário.

Já a ameaça *CL Completed* ou *"Cancel Location Completed*" se demonstra quando uma requisição enviada por um nó não autorizado é recebida pela rede, solicitando a exclusão do usuário no banco de dados temporário, VLR (*Visitor Location Register*), sendo uma ameaça da classe "Negação de Serviço" do usuário. Por fim, a ameaça *PSI Completed* ou "*Provider Subscriber Information Completed"* é caracterizada como uma solicitação não-autorizada de obtenção dos dados de localização do assinante para o banco de dados temporário. Essa solicitação por padrão é atendida pela rede e pode ser identificada como parte das ameaças de "Interceptação".

3.2.3 Distribuição da origem das ameaças

A fim de verificar se uma única origem era responsável pelas ameaças ou se uma ampla gama de fontes geravam as mesmas, foi feita uma observação mais específica para cada ameaça principal mencionada na subseção anterior. Na Figura 3.6, tem-se a quantidade de atacantes para cada ameaça descrita. Desta forma, é possível identificar quais ameaças são as mais distribuídas e complexas de reproduzir. Algumas ameaças, como o *Abnormal TCAP handshake Aborted* e o *Provider Subscriber Info* possuem poucas fontes. Porém, estes produzem ameaças em grandes quantidades, enquanto no *Fast Relocation*, a quantidade massiva de ameaças pode ser justificada pela quantidade de origens. A ameaça *Fast Relocation with SMS* só começou a ser originada a partir do segundo mês de observação. Já a ameaça *Abnormal TCAP handshake Aborted* teve duração de apenas alguns dias, porém significativos para esta ameaça estar na vice-liderança do ranqueamento mesmo com uma pequena quantidade de atacantes.

Por meio de um identificador, é possível verificar quais localidades são originárias das ameaças assim como seus principais alvos, permitindo inclusive reconhecer quais operadoras eram ameaçadas através dos 5 primeiros números do IMSI alvo. A Tabela 3.1 mostra um resumo com a distribuição geográfica para atacantes em redes móveis, bem como os principais países-alvo. Ao todo, 158 países disparam ameaças SS7, advindas de diferentes fontes, com liderança dos EUA, responsáveis por 32,77%



(a) *Fast Relocation.* (b) *Abnormal TCAP handshake Aborted.*

(c) *Fast Relocation* com SMS. (d) *SendRoutingInfoForSM.*

(e) *CL Completed.* (f) *Provider Subscriber Info.*

Figura 3.6: Número de fontes atacantes simultâneas por tipo de ameaça.

do total de ameaças enviadas. Em contrapartida, 58 países são atingidos, sendo o Brasil o principal alvo, concentrando 96,67% da quantidade total de ameaças recebidas. É importante citar que para o levantamento deste cenário, foi realizada uma filtragem e apenas os MNC (*Mobile Network Code)* dos provedores de serviço declarados como possíveis alvos foram avaliados.

3.2.4 Duração das ameaças

A Figura 3.7 mostra as distribuições empíricas da duração das ameaças. Devido ao grande volume de dados, uma amostra foi usada para geração destas distribuições, compreendendo apenas os meses de Dezembro e Janeiro. Observa-se que 99,7% das ameaças gerais que ocorrem (classificadas como "Total") tem a duração de até 10

Tabela 3.1: Estatísticas das principais origens e destinos das ameaças.

Posição	**País**	**Fontes distintas**	**%**	**País**	**Alvos distintos**	**%**
1	Estados Unidos	133	32,77	Brasil	18	96,67
2	Chile	34	27,73	França	1	0,94
3	Reino Unido	99	4,13	Paraguai	1	0,67
4	Itália	93	3,94	Uruguai	1	0,42
5	Cambodja	1	2,81	Porto Rico	1	0,38
	outros	181	28,62	outros	27	0,92
Total	*158 países*	541	100,0	*58 países*	50	100,0

minutos e 80% das ameaças para esta categoria acontecem quase que simultaneamente, o que mostra que a frequência de ameaças é alta e o intervalo entre duas ameaças consecutivas é pequeno. A média total é de 1 ameaça a cada 38 segundos. Já para a ameaça *Cancel Location*, a distribuição da duração das ameaças é mais uniforme, com duração até 60 minutos, apresentando uma menor frequência de ameaças. Apesar das ocorrências terem sido esparsas para *Cancel Location*, 78% das ameaças deste tipo ainda são disparadas de forma quase simultânea. A operadora pode, como contramedida, adotar o critério de filtragem após o recebimento de um determinado número de ameaças iguais, que podem ser consideradas como uma tentativa de negação de serviço, por exemplo, pela característica de rajada.

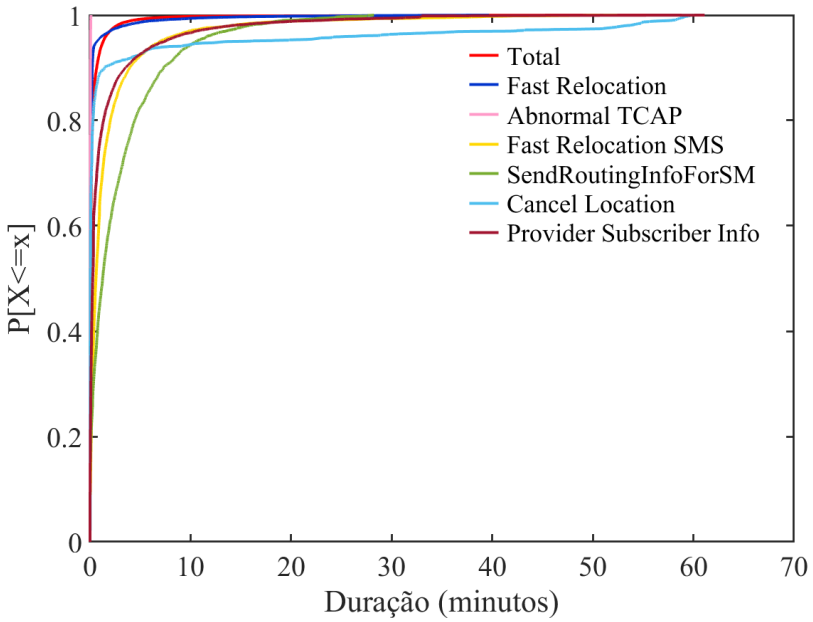


Figura 3.7: CDF dos meses de Dezembro/2018 e Janeiro/2019.

Tabela 3.2: Probabilidade empírica de transições entre ameaças consecutivas.

Ameaças	**Fast Relocation**	**Abnormal TCAP**	**FR with SMS**	**Send Routing Info**	**CL Completed**	**Provide Sub. Info**	**Others**
Fast Relocation	0,8690	0,0006	0	0,0113	0,0263	0,0076	0,0853
Abnormal TCAP	0,0024	0,9758	0	0,0006	0,0079	0,0002	0,0130
FR com SMS	0	0	0,0625	0,2500	0,1250	0	0,5625
Send Routing Info	0,0440	0,0005	0,0007	0,6698	0,0531	0,0179	0,2139
CL Completed	0,2061	0,0146	0,0004	0,1188	0,392	0,0332	0,2978
Provides Sub. Info	0,0789	0,0005	0	0,0356	0,0392	0,6817	0,1641
Outros	0,1736	0,0059	0,0010	0,1178	0,0841	0,0330	0,5846

3.2.5 Transições entre ameaças consecutivas

Para observar o comportamento da rede após o recebimento de um determinado tipo de ameaça e quais as decorrências da mesma, foram calculadas as probabilidades de transições entre as ameaças mais numerosas. A Tabela 3.2 resume os resultados. Cada elemento π_{mn} da tabela representa o total das transições entre as ameaças m e n consecutivas (τ_{mn}) sobre o total de transições observadas. Considerando que $m, n \in \mathcal{A}$, onde $\mathcal{A}$ é o conjunto de ameaças analisadas, $\pi_{mn} = \tau_{mn}/(\sum_{i\in\mathcal{A}} \sum_{j\in\mathcal{A}} \tau_{ij})$. A ameaça m e o n são, respectivamente, os títulos da linha e da coluna da tabela. Por exemplo, após a ocorrência de uma ameaça *Provide Subscriber Info*, a possibilidade de ocorrer uma ameaça do tipo *Fast Realocation with SMS* não é observada, já que o valor desta transição na Tabela 3.2 é zero. Para todos os casos, pode-se verificar que após uma ameaça ser executada, a maior probabilidade é que em sequência seja disparada o mesmo tipo de ameaça anterior, como pode ser observado pelos maiores valores encontrados na diagonal principal da tabela. A maior probabilidade da ocorrências de ameaças repetidas indica a presença de ameaças em rajadas, o que é uma importante característica a ser explorada pelas contramedidas a serem empregadas pelas operadoras.

Os resultados permitiram atestar que a operadora estava exposta a ameaças em SS7. Após a detecção das ameaças, a operadora tomou algumas providências, como o bloqueio de todas as mensagens da Categoria 1 eleitas pelo GSMA, como a *AnyTimeInterrogation*, conforme a IR.82 [30]. Para as demais categorias, uma verificação mais apurada deve ser feita, realizando a checagem da origem do assinante (interno ou externo) e comparando parâmetros para atestar a legitimidade da mensagem. Portanto, ainda não há uma solução definitiva para ameaças destas categorias.

Com o objetivo de desenvolver uma solução efetiva de mitigação e como forma

de implementação de um ciclo de segurança de rede, o primeiro passo é a avaliação da exposição da rede, conforme apresentado neste Capítulo. O passo seguinte é a auditoria da rede, que permita o total conhecimento das operações de rede. Tais informações, como por exemplo, a identificação dos atacantes, o desencadeamento das ameaças na rede e a análise de como as ameaças estão interconectadas, devem ser armazenadas e estar disponíveis para verificação. A corrente de blocos, por suas propriedades intrínsecas, como imutabilidade, confiabilidade e auditabilidade, se mostra uma proposta interessante para atendimento a esta demanda de auditoria.

Capítulo 4

Proposta de Auditoria para Redes Móveis

Este capítulo propõe uma forma de auditoria das movimentações da rede com foco principal no monitoramento das atividades de possíveis atacantes, sendo um complemento às contramedidas existentes. Nesse contexto, a introdução da tecnologia corrente de blocos é proposta, já que permite a verificação de todas as alterações que são feitas na rede e, consequentemente, a identificação dos atacantes e a frequência com que a rede é ameaçada. Tais informações permitem mapear a rede e desenvolver estratégias que mitiguem as ameaças. Para isto, primeiramente será revista a fundamentação teórica de corrente de blocos, como suas características, sua estrutura, suas categorias e algoritmos de consenso. A seguir, será apresentada uma proposta de auditoria aplicando a corrente de blocos no problema de vulenrabilidades em SS7 para as redes móveis. Em sequência, será feita uma avaliação da proposta por meio do desenvolvimento de um contrato inteligente e testes de desempenho. Por fim, são apresentados os resultados dos testes de desempenho e os trabalhos relacionados aos tópicos abordados.

4.1 Corrente de Blocos

A corrente de blocos é uma tecnologia revolucionária que permite mudar a forma de relação entre duas partes. Consiste em um livro-razão (*ledger*) distribuído e imutável que registra diferentes tipos de transações de forma permanente sem a necessidade de uma entidade central comum ou um nó intermediário [33]. Isso permite que duas partes que não confiem entre si interajam de forma segura, sem necessidade de um terceiro confiável. Ademais, possibilita a descentralização em uma rede par a par. Foi popularizada através do uso das criptomoedas *Bitcoin* e hoje é utilizada para as mais diversas aplicações além das transferências financeiras,

como aplicações em IoT [34–36], saúde [37–39], agricultura [40, 41], dentre outras.

Consiste em uma estrutura de dados composta por uma corrente de blocos, onde cada bloco é ligado ao antecessor por um resumo criptográfico (*hash*), formando a corrente. Todo bloco possui um conjunto de transações inseridas de forma ordenada, sendo estas previamente validadas. As transações representam operações atômicas para transferência entre as partes. Os nós que emitem as transações e posteriormente geram o bloco, são chamados de nós mineradores.

As decisões para submissão dos blocos à corrente são feitas por meio de algoritmos de consenso, de forma descentralizada, onde os nós da rede devem chegar a um acordo sobre a adição do bloco à corrente. Uma vez aceito, o estado global da corrente é atualizado e o protocolo de consenso replica a informação atualizada da corrente para todos os nós participantes da rede. Como principais atributos da corrente de blocos [42, 43], pode-se citar:

- **Descentralização:** A transferência de ativos entre duas partes não depende de uma entidade confiável centralizadora, a troca de transações é feita em uma rede par a par, sem necessidade de intermediários. As decisões realizadas no sistema são feitas de forma descentralizada através de algoritmos de consenso.

- **Imutabilidade:** Os dados armazenados na corrente de blocos não podem ser alterados visto que cada bloco é ligado com o anterior pelo resumo criptográfico, o que configura a imutabilidade e integridade da corrente.

- **Auditabilidade:** Todos os dados armazenados na corrente de blocos estão visíveis a todos os nós participantes (propriedade de transparência). Isto implica em disponibilidade de verificação dos dados mantidos na corrente através do acesso a qualquer registro publicado, permitindo a rastreabilidade das transações e a possível identificação da ação de atacantes bem como a análise do desencadeamento das ações maliciosas.

- **Redundância:** Os dados são distribuídos entre todos os nós participantes da rede par a par, o que garante redundância das informações e disponibilidade do sistema mesmo em meio a falhas.

Estas características são importantes em uma rede vulnerável a atacantes, pois como vimos no Capítulo 2, cada ameaça fere alguma premissa de segurança de rede, tais como: confiabilidade, integridade dos dados e disponibilidade dos elementos de rede [29]. O modelo de corrente de blocos, por sua vez, possui atributos que atendem à estas necessidades e confere maior robustez ao sistema, podendo ser uma alternativa interessante para as redes expostas.

4.1.1 Tipos de Correntes de Blocos

A corrente de blocos pode ser classificada como pública, privada, permissionada ou não permissionada. A divisão entre pública ou privada refere-se a acessibilidade dos nós à corrente. Em uma corrente de blocos pública, qualquer nó pode ingressar e sair da rede a qualquer momento. Não existe uma restrição para adesão à corrente nem quanto a visibilidade dos dados, que são expostos na Internet. A criptomoeda *Bitcoin*, por exemplo, adota este modelo [33].

Já no caso de uma rede privada, apenas nós autorizados têm acesso à rede, o que garante uma maior confidencialidade de informações. Tem característica centralizada, geralmente são governadas por instituições e, portanto, possuem maior eficiência e menor confiabilidade.

A classificação entre redes par a par permissionadas e não permissionadas é concernente à atuação dos nós na rede [42]. Em redes não-permissionadas, todos os nós desempenham igualmente os mesmos papéis, sendo conjuntamente responsáveis pela geração e validação de transações, construção de blocos e participação de consenso. Ao contrário destas, em redes permissionadas, os nós executam papéis distintos quanto a participação do algoritmo de consenso e a geração de blocos.

4.1.2 Algoritmos de Consenso

Os mecanismos de consenso proveem a forma descentralizada de tomada de decisões no sistema da corrente de blocos. O algoritmo de consenso baseado em prova mais conhecido é a Prova de Trabalho (*Proof of Work - PoW*), utilizado pelo *Bitcoin* e em redes públicas [33]. A submissão de um bloco na corrente é feita através da resolução de um desafio matemático pelos nós mineradores da rede. O nó que consegue resolver o desafio recebe incentivos em troca do trabalho. Esse tipo de algoritmo, porém, possui grande custo computacional e latência para convergência da rede, o que compromete a escalabilidade [43]. Outro algoritmo utilizado em redes par a par públicas é a Prova de Posse (*Proof of Stake - PoS*). É utilizado pela criptomoeda *Ethereum*, onde a troca de transações é feita por meio de contratos inteligentes. Diferentemente da Prova de Trabalho, a mineração é feita com base em investimentos de participação dos nós, em forma de apostas. Portanto, quanto maior a riqueza de um nó, maior a chance de submissão de um bloco.

Em redes privadas, foi proposta a Prova de Autoridade (*Proof of Authority - PoA*) [44], usado no *Parity Ethereum* [45], no qual um conjunto de nós previamente definidos possuem autoridade para participar do consenso. O PoA assume que esses nós são confiáveis [46]. O esquema de mineração é o de rotação (chamado *mining rotation*), onde cada nó pode propor um bloco a cada rodada.

Há ainda os protocolos de consenso determinísticos, onde as tomadas de decisões

de rede são baseadas em votação. A aprovação dos blocos se dá pela decisão de um quórum, portanto presume-se que os mineradores são poucos e confiáveis. Ademais, após a decisão dos mineradores, o novo bloco é divulgado na rede de forma síncrona e consistente de modo que todos os nós tenham a mesma visão da rede. Esse modelo também é aplicável para redes privadas e pode-se destacar o algoritmo Prático de Tolerância a Falhas Bizantinas (*Practical Byzantine Fault Tolerance - PBFT*). O PBFT considera que os nós podem ser mal intencionados ou podem apresentar falhas, comprometendo a entrega das mensagens trocadas entre os participantes. O PBFT apresenta menor custo de processamento que o *PoW* e alta consistência, porém a custo do aumento da complexidade de mensagens, podendo comprometer o desempenho da rede [44, 47].

Dada a revisão da fundamentação teórica da corrente de blocos, a seguir, é proposto o uso dessa tecnologia em um ambiente não confiável, como o de redes móveis com uso do SS7.

4.2 Auditoria de Redes Móveis usando Correntes de Blocos

Com o intuito de verificar a aplicabilidade da solução de corrente de blocos no problema de vulnerabilidade da rede móvel e mensurar a vazão de transações emitidas na corrente de blocos para este caso, foi feita a verificação da quantidade de mensagens MAP trocadas na rede. Para isto, no mesmo cenário descrito no Capítulo 3, foi observado um traço de 1 hora de duração, cujos dados foram obtidos através do programa *tshark*. Foi desenvolvido um *script* em *AWK* para extração e contagem das mensagens MAP. A Figura 4.1 exibe o resultado da análise. No pico de mensagens, a quantidade de mensagens atingiu 317 requisições MAP/min, o que em segundos, chegaria a aproximadamente 6 requisições MAP/s. Esta quantidade é atendida tanto por algoritmos de consenso probabilísticos (por exemplo, o PoW, que permite o processamento de 7 transações/s [48]) quanto determinísticos, como os protocolos BFT, que possuem uma escalabilidade muito maior devido à pouca quantidade de nós [49].

As ameaças de personificação que foram abordadas neste trabalho consistem na intrusão de elementos falsos na rede SS7, ou seja, um atacante que consegue simular um elemento legítimo de rede e, assim, obter informações sensíveis dos usuários ou da própria rede. A tecnologia de corrente de blocos é interessante para este problema visto que confere propriedades de auditabilidade e transparência de rede, permitindo a implementação de um ciclo de segurança de rede, onde o primeiro passo é a auditoria, seguida de detecção e mitigação.

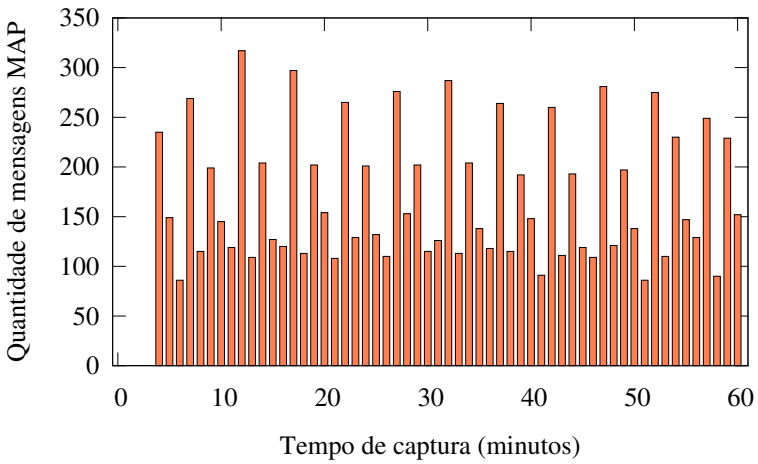


Figura 4.1: Quantidade de mensagens MAP recebidas na rede da operadora no intervalo de uma hora.

Trazendo os conceitos expostos sobre correntes de blocos ao contexto de redes móveis, cada mensagem MAP representa uma *transação*. Portanto, cada pedido MAP (*Request* ou *Response*) independente da fonte, é considerado uma transação. Com o objetivo de promover uma auditoria na rede que permita identificar as ameaças e garantir rastreabilidade das movimentações de rede, idealmente todas as transações MAP geradas são aceitas para que seja construído um histórico permanente das operações realizadas na rede. Algumas condições de validação, porém, são estabelecidas, tais como: uma transação somente é aceita caso obedeça aos requisitos mandatórios da especificação do 3GPP para o protocolo MAP (isto é, caso seja legítima) e ignorada quando não atendê-los. Sendo assim, se a semântica da mensagem MAP recebida estiver incorreta de acordo com a definição do 3GPP, esta é considerada inválida e consequentemente descartada. Caso haja alguma transação vazia, sem nenhum conteúdo, a mesma também será descartada. A validação de cada transação é feita localmente em cada nó pertencente ao algoritmo de consenso.

Como pretende-se aplicar a corrente de blocos a redes corporativas, governadas por uma instituição, a escolha por redes privadas é a mais adequada. Uma rede de operadora pode ser caracterizada como privada e distribuída. Dessa forma, o modelo de corrente de blocos que mais se aplica é o de redes privadas. Ademais, sobre a classificação de permissão, apesar da rede ser privada, esta é suscetível a ameaças externas, já que ainda não existem ferramentas de bloqueio efetivas a mensagens das categorias 2 e 3 do GSMA [30]. Isso significa que qualquer indivíduo com recursos necessários pode enviar mensagens MAP verdadeiras à rede, desde que atenda às especificações do 3GPP. Sendo assim, um elemento de rede que emita uma transação legítima torna-se um nó que atende os requisitos da rede. Isto posto, a rede é permissionada, na qual os nós executam papéis distintos. Apenas nós confiáveis (pertencentes à rede da operadora) podem participar dos processos inerentes à tec-

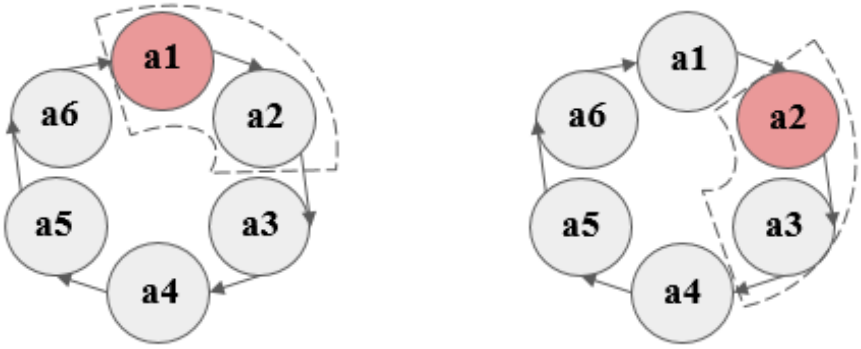
nologia corrente de blocos na rede, como validação de transações, geração de blocos e algoritmo de consenso. Como a finalidade é a auditoria da rede, é necessário que o algoritmo de consenso aceite o máximo de transações (idealmente todas), sem descartá-las, para manutenção de um histórico completo das transações submetidas à inserção na corrente de blocos. Com relação ao algoritmo de consenso, o escolhido foi o PoA, com uma implementação chamada *Clique*, por se tratar de um algoritmo para correntes de blocos privadas, pela simplicidade na submissão de blocos e pelo baixo fluxo de mensagens trocadas na rede. O funcionamento do algoritmo de consenso é explicado em detalhes na próxima subseção.

4.2.1 Algoritmo de Consenso Utilizado

Os algoritmos de consenso aplicáveis a redes públicas, como o PoW e o PoS foram descartados para aplicabilidade no prolema, visto que o modelo adotado é o de redes privadas.

Angelis *et al.* fazem uma comparação entre os algoritmos de consenso PBFT e PoA através do teorema CAP (*Consistency, Availability, Partition Tolerance*) [44], onde Consistência é definida pela igualdade de informações contidas em um nó, ou seja, todos possuem a mesma visão da rede; Disponibilidade refere-se à operacionalidade do sistema e por fim, Tolerância à Partição, que é a capacidade do sistema se manter operacional ainda que ocorram falhas que dividam as conexões do sistema. Os autores afirmam que é impossível um sistema distribuído apresentar plenamente as três propriedades. Ademais, foi feita uma análise com relação ao desempenho dos dois protocolos, em termos de tempo de convergência. Na avaliação, em termos de consistência, o PBFT se mostrou melhor do que o PoA, já que a versão "*Clique*"do algoritmo (explicada em detalhes mais adiante) pode possuir bifurcações. Além disso, por se tratar de um algoritmo baseado em prova, o PoA é um protocolo probabilístico, onde há uma sincronia eventual, garantindo que os nós irão receber atualizações das informações da corrente de blocos eventualmente, comprometendo a consistência do sistema. O PBFT por sua vez, é um modelo determinístico, onde após a decisão feita pelo quórum, todos os nós são atualizados de maneira síncrona. Com relação às outras propriedades, como disponibilidade, tolerância a partições e desempenho, o *Clique* (PoA) se mostrou uma alternativa melhor, pois o PBFT possui mais rodadas para chegar a um consenso e também gera mais mensagens à rede, apresentando maior complexidade para execução. Portanto, para este trabalho, foi priorizado o desempenho já que para os elementos da rede móvel (onde pode-se incorporar a funcionalidade de execução de uma corrente de blocos) não é desejável o aumento da complexidade de mensagens e nem a indisponibilidade do sistema de consenso, sendo assim, o protocolo escolhido foi o PoA.

O PoA atribui autoridade aos nós e baseia-se em períodos definidos. A cada rodada, um subconjunto dos nós de autoridade pode propor um bloco, sendo um desses nós o líder da rodada. Este subconjunto é composto por no máximo $N - (\frac{N}{2} + 1)$ autoridades, sendo N o número total de nós de autoridade na rede [44]. Por exemplo, caso existam seis nós de autoridade, apenas dois podem enviar blocos a cada intervalo de tempo, sendo o primeiro deles o líder. Ademais, uma autoridade deve submeter um bloco a cada $\frac{N}{2} + 1$ blocos (para evitar falhas Bizantinas). A Figura 4.2 ilustra o esquema utilizado no algoritmo de consenso PoA, com seis nós de autoridade e apenas dois habilitados a enviar blocos para a corrente num dado instante de tempo. A Figura 4.2(a) representa a primeira rodada no tempo $t1$, enquanto a Figura 4.2(b) representa o próximo intervalo $t2$, apresentando o esquema de rotação do algoritmo de forma cíclica, que distribui de forma justa a responsabilidade de criação dos blocos.



(a) (b)

Figura 4.2: Esquema cíclico de mineração do algoritmo PoA, composto pelo conjunto dos nós de autoridade. Dentro do pontilhado, os nós que podem enviar um bloco a cada instante, sendo o líder da rodada representado em rosa. (a) No instante $t1$, o $a1$ é o líder; (b) enquanto no instante $t2$, o $a2$ é o líder.

A partir do início de uma nova rodada, o líder atual deve propor um bloco, ainda que este não possua nenhuma transação, i.e., propor um bloco mesmo que esteja vazio. Os nós não-líderes podem propor um bloco quando estiverem habilitados para fazê-lo, dentro de sua rodada, após aguardarem um tempo aleatório. A Figura 4.3 ilustra o funcionamento do algoritmo de consenso PoA. O envio de um bloco pelos nós habilitados é feito em *broadcast* para todo o conjunto de nós de autoridade. Cada bloco proposto pelo líder é diretamente inserido na corrente de blocos em todos os nós participantes. Porém, neste trabalho, adota-se como premissa que blocos sem nenhuma transação são imediatamente descartados. O bloco emitido pelo líder tem um peso maior com relação ao bloco de um nó comum, para evitar bifurcação na corrente.

A Figura 4.4 apresenta a arquitetura proposta baseada no modelo PoA. Cada

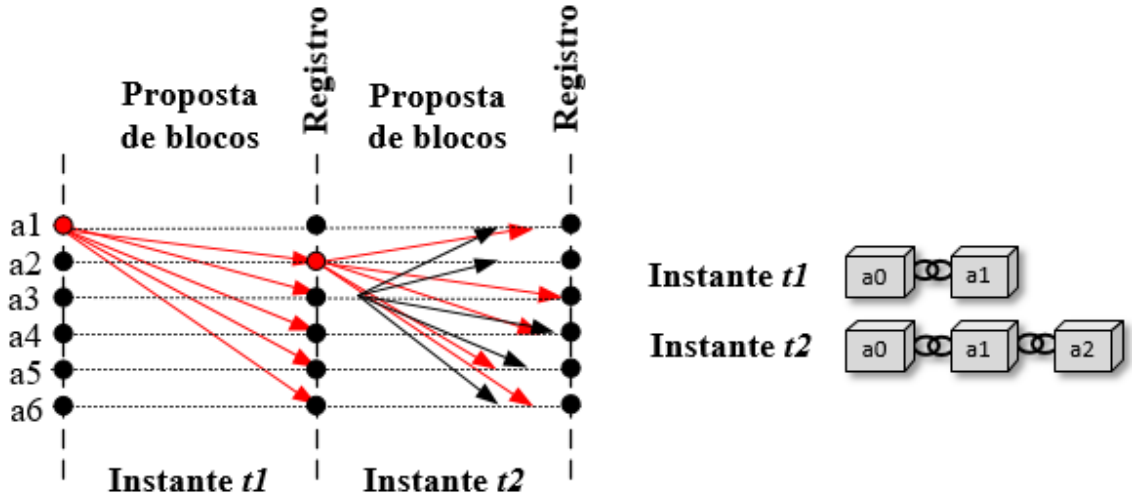


Figura 4.3: Funcionamento do algoritmo de consenso PoA. O líder a1 propõe um bloco no instante $t1$ e este é inserido na corrente. Após isso, em uma nova rodada, o líder a2 propõe um novo bloco, porém o nó a3 também faz sua proposta. Para os nós a4 e a5, o bloco a3 chegou primeiro. Este problema de bifurcação é resolvido com os pesos dados aos blocos gerados pelo líder.

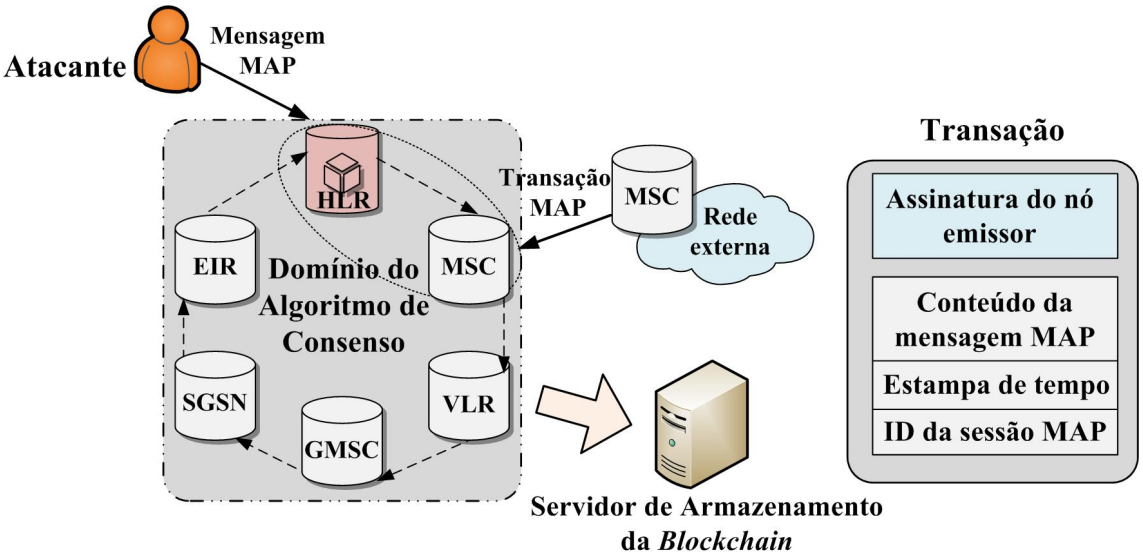


Figura 4.4: Arquitetura proposta com base no algoritmo PoA. Um modelo de transação também é proposto, contendo um identificador que possibilita o rastreio das ações desencadeadas pela transação inicial.

elemento de núcleo da rede móvel da operadora (vide Figura 2.1) compõe o conjunto de nós de autoridade, que são os nós confiáveis que participam efetivamente dos processos executados na formação da corrente de blocos. O conteúdo da transação proposta é constituído pela mensagem MAP (pedido ou resposta), por uma estampa de tempo, assinatura do nó emissor da transação (o primeiro nó que recebeu a requisição) e um ID que caracteriza a sessão na qual essa transação é pertencente. Entende-se como sessão o conjunto de todas as operações geradas decorrentes de

uma transação inicial, percorrendo o caminho desde a sua entrada na rede até o retorno para o remetente da transação. Na Figura 2.7, a sessão seria definida desde a mensagem ATI até a sua resposta final.

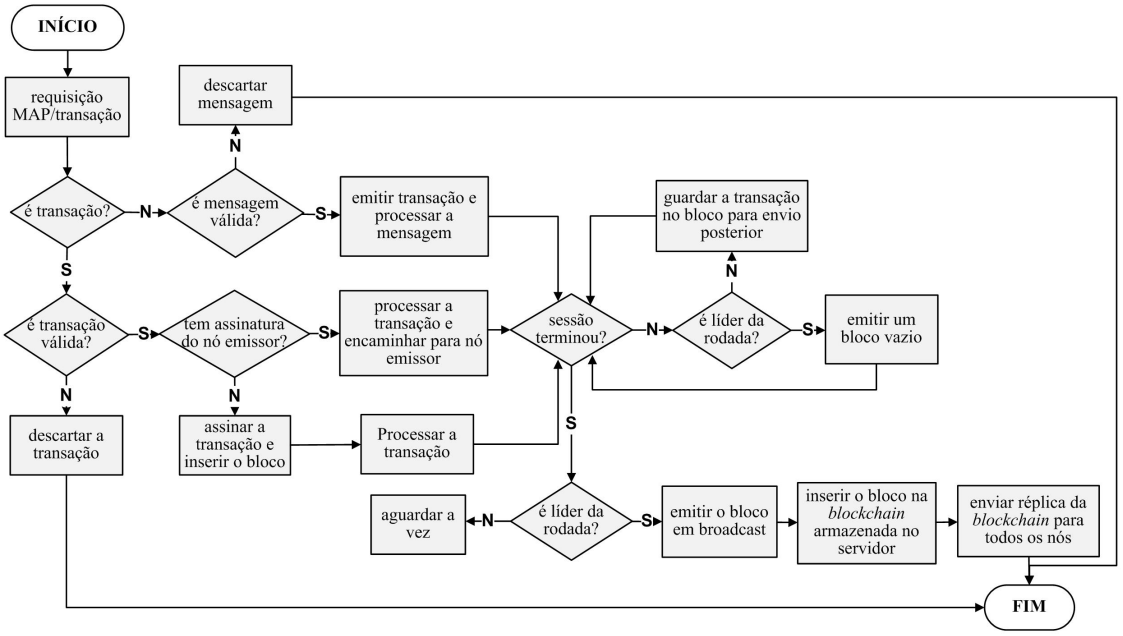


Figura 4.5: Fluxograma operacional para emissão de um bloco e inserção na corrente de blocos proposta.

A Figura 4.5 mostra o fluxo de operações a serem executadas para inserção de um bloco na corrente de blocos. Quando a rede receber uma mensagem MAP, esta é validada localmente pelo nó que a recebeu. Após a validação, este nó encapsula a mensagem em uma transação e a assina, a fim de que todas as mensagens relativas àquela sessão sejam encaminhadas a ele para montagem do bloco, demonstrando a sua responsabilidade sobre a emissão do mesmo. Feito isso, o nó processa a transação, executando na rede a ação correspondente à requisição recebida. Cada nova transação resultante deste pedido é identificada com o mesmo valor (ID) de sessão. Paralelamente, os nós executam o algoritmo PoA. Na Figura 4.4, por exemplo, o HLR é o líder da rodada e precisa propor um bloco. O bloco será formado por todas as transações pertencentes à mesma sessão, organizadas cronologicamente pela estampa de tempo. O líder da rodada somente envia um bloco quando possuir o conjunto total de transações da mesma sessão. Cada novo bloco inserido na corrente é armazenado no servidor de armazenamento da corrente de blocos. Este servidor é opcional, sendo utilizado apenas se a rede não possuir capacidade de armazenamento em seus nós. Cada nó de autoridade possui uma réplica da corrente de blocos divulgada a todos a cada nova rodada do algoritmo de consenso. Deste modo, todos os nós enxergam a corrente de blocos da mesma forma.

A Figura 4.6 mostra a corrente de blocos gerada pelo processo. Note que as

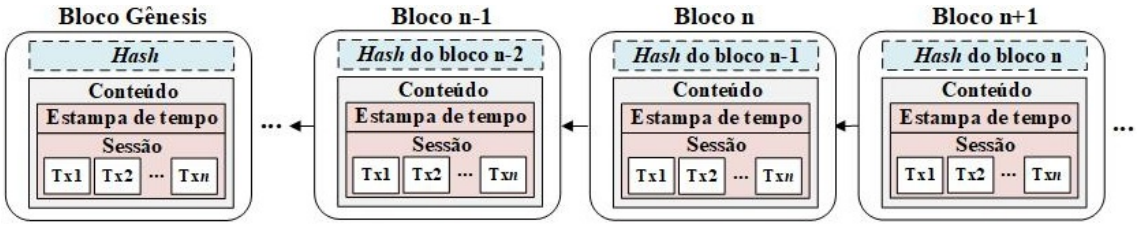


Figura 4.6: A estrutura da corrente de blocos proposta. As transações armazenadas em cada bloco são relacionadas entre si pela ID de sessão.

transações, Tx1 a Txn, são pertencentes à mesma sessão e compõem o conteúdo do bloco. A corrente de blocos gerada, portanto, registra de forma permanente todas as operações trafegadas pela rede, possibilitando a auditoria pela verificação da corrente de blocos.

4.3 Avaliação da Proposta

Para avaliação da proposta, foi desenvolvido um contrato inteligente na plataforma *Hyperledger Fabric*. As transações são definidas em contratos inteligentes no HLF. O código apresenta as seguintes funções: o cadastro de um assinante na rede, a consulta de informações de um assinante já registrado e a remoção de um assinante do banco de dados, onde cada função descrita é uma transação a ser executada na rede. Esses tipos de transações são as mais frequentes realizadas pelos atacantes, conforme detalhado no Capítulo 2. Uma vez que o atacante possua o número de telefone do assinante e acesso à rede SS7, o mesmo efetua consultas legítimas à rede para obtenção de outras informações a fim de ocasionar um ataque de maior proporção, podendo inclusive remover o assinante da rede. Além destas, também foi desenvolvida uma função para emissão de uma transação, com o objetivo de executar a proposta na corrente de blocos, contendo a assinatura do nó originador das transações MAP (inserção, consulta ou remoção), uma estampa de tempo e um ID da transação, que caracteriza a sessão à qual a mesma pertence.

Na próxima subseção, o *Hyperledger Fabric* é apresentado, bem como o algoritmo desenvolvido e a topologia utilizada. Após isso, são feitos testes de desempenho através do *Hyperledger Caliper*, que exibe como resultados a vazão (em transações/segundo ou *tps*), a latência e o consumo de CPU para cada transação.

4.3.1 *Hyperlegder Fabric*

O *Hyperledger Fabric*[1] (HLF) é um ambiente de desenvolvimento para correntes de blocos baseadas em redes par a par privadas e permissionadas [50]. O HLF é um

[1]https://www.hyperledger.org/projects/fabric

dos ambientes que fazem parte de um projeto maior chamado *Hyperledger*, que é uma iniciativa mantida pela *Linux Foundation*, criada para alavancar a tecnologia de corrente de blocos entre indústrias. O *Hyperledger* é uma plataforma de colaboração global e de código aberto que possui mantenedores em várias áreas da indústria, tais como finanças, tecnologia, suprimentos, etc. Além disso, possui várias plataformas, dentre as quais destaca-se o *Fabric*, por sua maturidade de implementação e popularidade em desenvolvimento voltado para empresas. O HLF é uma plataforma de livro-razão distribuído que possui uma arquitetura modular e flexível, permitindo a customização de ambientes de correntes de blocos. O HLF permite ainda a criação de contratos inteligentes em diferentes tipos de linguagem de programação (*NodeJS, Golang e Java*) e usa como algoritmo de consenso o CFT *(Crash Fault Tolerance)*.

A rede do HLF é composta por nós de diferentes funções, pertencentes à organizações. Podem ser de 3 tipos, dependendo do papel desempenhado: Cliente, Pares (*Peers*) e Ordenadores [50]. O nó Cliente é responsável por emitir as transações, assinando-as com uma assinatura digital, através do uso de uma chave privada. Os nós Pares validam e executam as transações solicitadas pelos clientes de acordo com a política criada pelos contratos inteligentes e também armazenam uma réplica da corrente de blocos. Os pares que instalam e instanciam os contratos inteligentes são chamados de "aprovadores"(*Endorsers* ou *Endorsing Peers*). Os nós Ordenadores são responsáveis pelo serviço de ordenação das transações que compõe um bloco, pela geração do bloco e ordenação na corrente através do algoritmo de consenso e pela divulgação dos blocos para os nós participantes da rede. Como se trata de uma rede par a par permissionada, todos os nós pares possuem uma identidade e pertencem à uma organização. A flexibilidade do HLF permite a criação de várias organizações, cada uma contendo um ou mais nós. A identidade utilizada na autenticação dos nós é atribuída pelo MSP *(Membership Service Provider)*.

Para comunicação entre os nós, é criado um canal, onde todos os participantes da rede são atribuídos. Os dados trocados entre a rede são unicamente transmitidos através deste canal, cuja comunicação é criptografada. É possível criar mais de um canal, dependendo do tipo de aplicação.

As transações são feitas por meio de contratos inteligentes, chamados *chaincodes*, os quais são instalados e instanciados nos nós pares. Tais contratos podem ser desenvolvidos em *Go*, *Nodejs* ou *Java*. A cada nova execução de um contrato inteligente ou de uma transação na rede, o estado global da rede é atualizado com a criação de um novo bloco que posteriormente é replicado para todos os nós da corrente de blocos.

A fim de avaliar a proposta, foi desenvolvido um contrato inteligente (*chaincode*) na linguagem *Go* no HLF. O código na íntegra encontra-se no Apêndice A. A função *"setTransaction"* estabelece alguns parâmetros para a transação a ser gerada, são

eles: a assinatura do nó emissor, o autor da transação; o ID de sessão que foi calculado como a assinatura criptográfica do nó emissor e a estampa de tempo. Além desta, o contrato possui outras 3 funções: *"setMAP"*, *"query"* e *"delete"*. O objetivo da *setMAP* é inserir um usuário na rede, simulando um processo de registro. Para fins de simplificação no cadastro, apenas o nome e o número do usuário estão sendo armazenados. Nesta função, são realizados dois testes: um para verificar se os dados são vazios, o que não é permitido para a transação, sendo o dado automaticamente descartado e a segunda para verificar se os dados seguem os critérios do 3GPP. Para isto, como exemplo, é feita uma verificação do comprimento do número informado, caso seja menor do que a quantidade estabelecida, a informação é descartada. Para o teste, foi estabelecido o valor de 13 dígitos, sendo 2 dígitos para o país (MCC), 2 dígitos para a operadora (MNC) e 9 dígitos para o número do assinante (MSISDN). Para as funções de consulta e remoção, basta informar o nome do usuário para executar a operação.

4.3.2 *Hyperledger Caliper*

O *Hyperledger Caliper*[2] (HC) é uma ferramenta que permite a avaliação de desempenho das aplicações de correntes de blocos dos demais projetos *Hyperledger* da *Linux Foundation* tais como o *Fabric*, *Composer*, *Sawtooth*, *Iroha*, etc. Os indicadores gerados pelo *Caliper* mostram resultados de vazão (transações por segundo), latência e uso de recursos computacionais (CPU e Memória).

Para execução dos testes de desempenho no *Caliper*, foi utilizada uma máquina virtual na nuvem oferecida pelo Google. A máquina possui como configuração 1vCPU com 4GB de memória RAM, processador Intel Xeon E5-2650 V3 (*Haswell*), 500GB de HD e sistema operacional Linux Ubuntu 16.04 com *kernel* versão 4.15.0-1040. A versão 1.4.0 do *Hyperledger Fabric* foi usada. Como topologia para os testes, foi adotado um modelo que possui 3 organizações e 2 nós pares pertencentes a cada uma, totalizando 6 nós (nomeados de *"peerX.OrgY.example.com"*, onde X é o número do par e Y é o número referente à organização. O *Caliper* executa em *docker* e cada nó está distribuído em um *docker* diferente. Além dos nós pares, há ainda o nó ordenador que faz a organização dos blocos dentro da corrente de blocos. Como modelo de base de dados, foi usado o *LevelDB*. Os testes foram feitos variando-se a quantidade de transações emitidas e o tamanho do bloco gerado pela corrente de blocos. O código do contrato inteligente foi importado ao *Caliper* e foram avaliadas as funções de inserção e consulta. Os dados gerados pela ferramenta são exportados em uma página HTML e foram posteriormente adicionados em uma planilha no Excel para geração dos gráficos com o MATLAB. Na próxima subseção,

[2]https://github.com/hyperledger/caliper

os resultados encontrados são discutidos.

4.3.3 Resultados Obtidos

O *Caliper* faz a simulação de uma rede de corrente de blocos emitindo várias transações por segundo, a fim de verificar alguns parâmetros como latência, vazão e consumo de recursos (memória e CPU). A Figura 4.7 mostra a latência encontrada nos testes, sendo que os valores apresentados referem-se à latência média encontrada. Latência no contexto da corrente de blocos é o tempo necessário para a plataforma em questão, *Hyperledger Fabric*, responder à uma transação. Pode-se observar que para a função de consulta, o tempo de resposta é praticamente constante e inferior ao tempo utilizado para cadastro de um usuário. Conforme esperado, quanto maior o número de transações a serem processadas, maior o tempo de latência. A Figura 4.8, exibe a vazão encontrada na avaliação, onde vazão corresponde ao número de transações bem sucedidas por segundo. Da mesma forma que a latência, para a função de consulta, a vazão se mostrou constante. Como a função *query* é apenas uma função de consulta, esta apresentou resultados melhores com relação à função *setMAP*, que exige maior processamento.

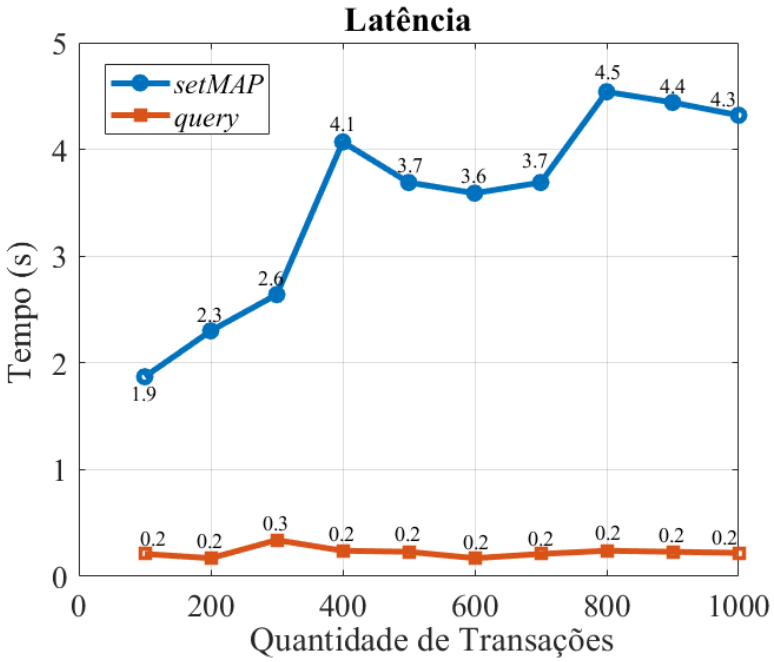


Figura 4.7: Latência encontrada nos testes de performance do *Hyperledger Fabric* para variação da quantidade de transações.

As Figuras 4.9 e 4.10 correspondem aos valores de latência e vazão com alteração do tamanho do bloco. Os tamanhos utilizados foram 99MB, 128MB, 256MB, 512MB, 1024MB e 2048MB. Nota-se que, para ambos os gráficos, não foi notada uma grande discrepância nos valores, permanecendo constante apesar da modificação do

tamanho do bloco utilizado.

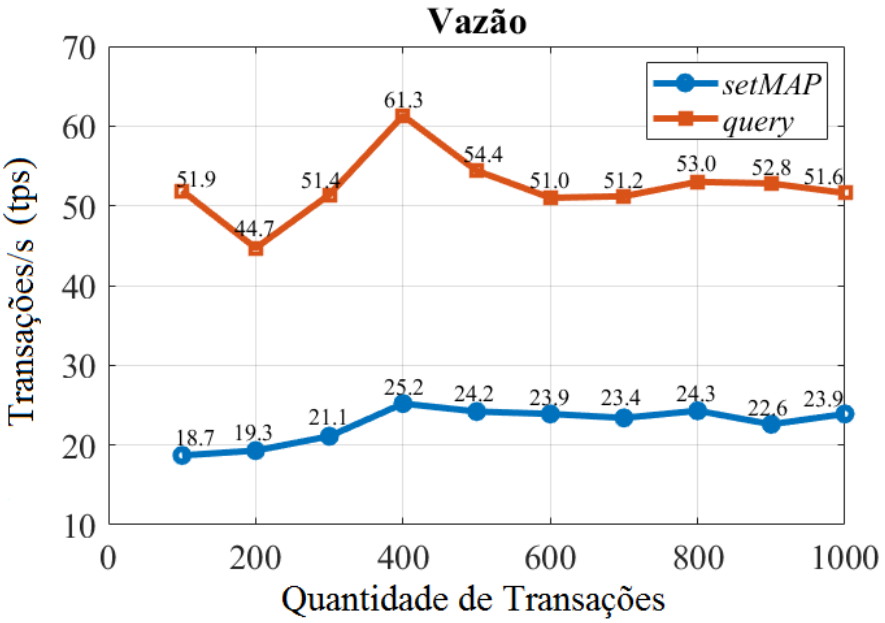


Figura 4.8: Vazão encontrada nos testes de performance do *Hyperledger Fabric* para variação da quantidade de transações.

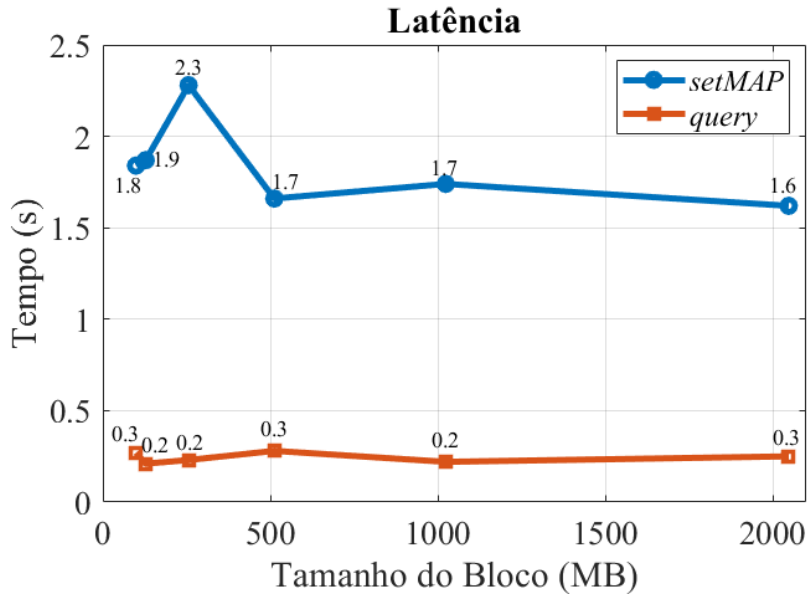


Figura 4.9: Latência encontrada nos testes de performance do HLF para variação do tamanho do bloco.

Além destes, foram observados os consumos dos recursos computacionais da máquina para a execução das transações do contrato inteligente. Todos os resultados para CPU e memória representam os valores máximos encontrados. No caso do consumo de memória, nota-se como esperado, que os pares consomem mais recursos do que o nó ordenador, por estes efetivamente executarem as transações. Isto

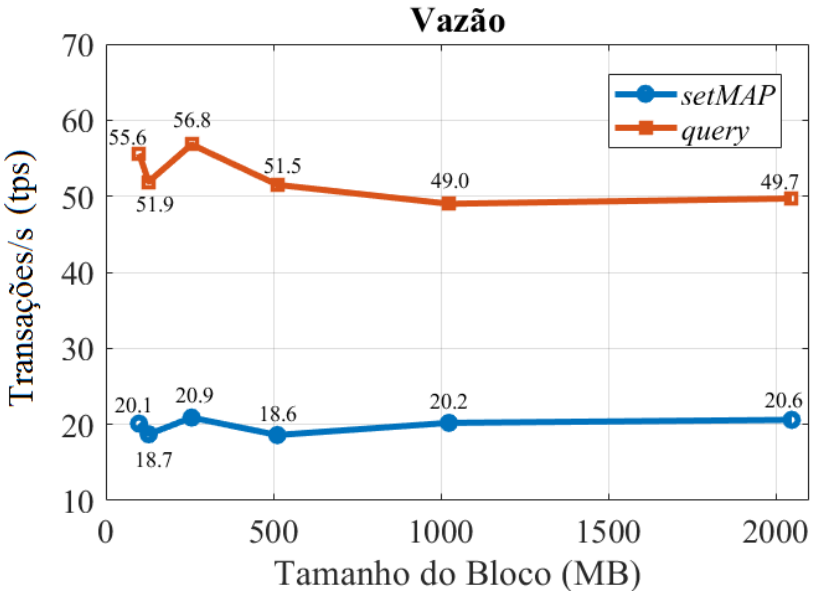


Figura 4.10: Vazão encontrada nos testes de performance do HLF para variação do tamanho do bloco.

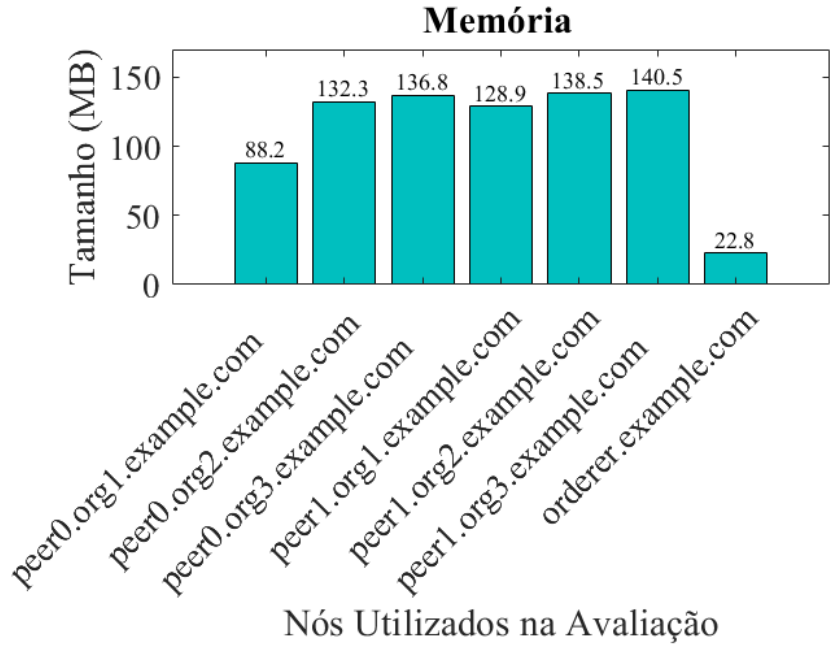


Figura 4.11: Consumo de memória nos pares utilizados nos testes de performance para 100 transações.

pode ser visto na Figura 4.11. Também foi alterado o tamanho do bloco para observar a variação do consumo de memória, mas esta tende a ser constante, vide Figura 4.12. No início do projeto, foi utilizada uma máquina de 4vCPUs, porém, esta estava superdimensionada, visto que com 1vCPU foi possível executar os testes sem problemas. Para o consumo de CPU, novamente os nós pares apresentaram maior valor, conforme mostrado na Figura 4.13. O tamanho do bloco também não impac-

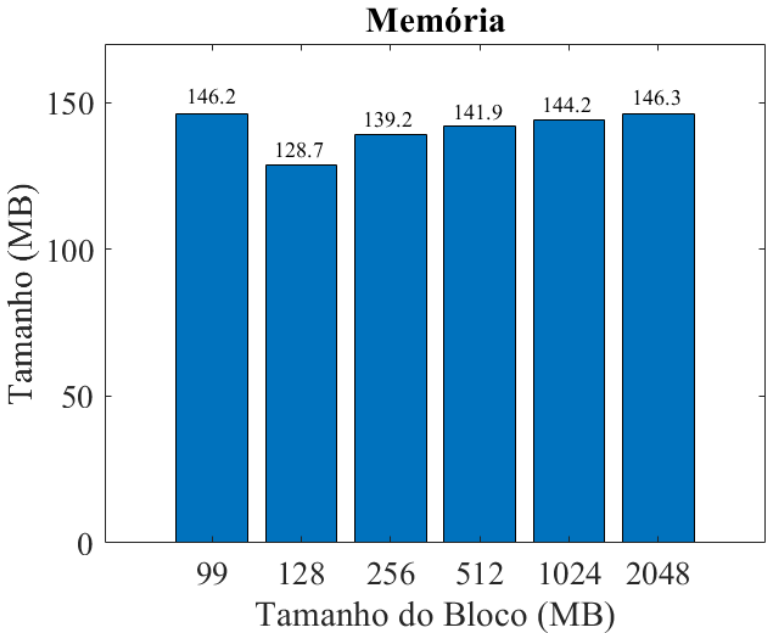


Figura 4.12: Consumo de memória nos testes de performance para variação do tamanho do bloco.

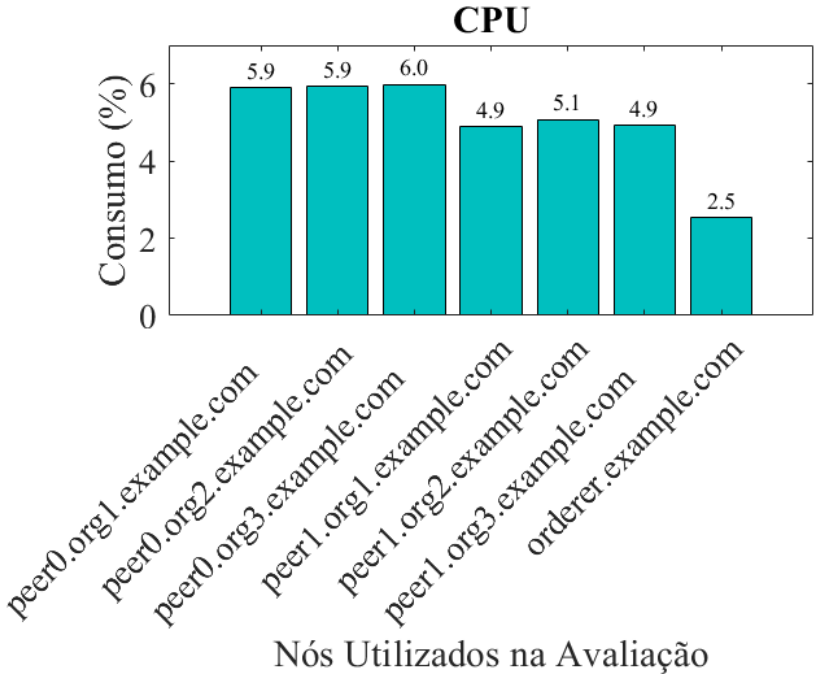


Figura 4.13: Consumo de CPU nos pares utilizados nos testes de performance para 100 transações.

tou significativamente no consumo de CPU, conforme pode ser visto na Figura 4.14. Isto pode ser explicado pelo fato de que independente de quantas transações caibam em um bloco, o que está sendo verificado é a capacidade do sistema em processar as transações. Com estes gráficos, pode-se concluir que a corrente de blocos é factível de ser utilizada em um ambiente real de uma provedora de serviços, visto que a

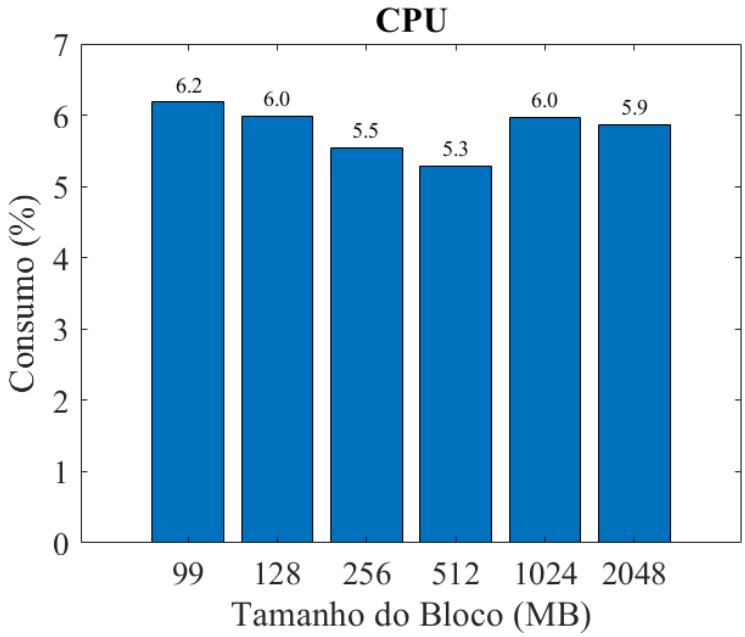


Figura 4.14: Consumo de CPU nos testes de performance para variação do tamanho do bloco.

capacidade utilizada na nuvem é bem semelhante ao que existe hoje em planta, onde os novos elementos já utilizam máquinas virtuais ao invés de uma solução proprietária. A corrente de blocos nesta configuração, não comprometeria, portanto, a execução de outros processos já existentes nos elementos de rede. A proposta deste trabalho não se baseia em apenas um único tipo de ameaça e também não possui como objetivo a mitigação das mesmas em primeiro momento. Entende-se que para o desenvolvimento de uma solução efetiva de mitigação, é necessário conhecer previamente o cenário dos ataques (tipos de ataques realizados, a relação entre os ataques e impacto na rede, a frequência dos ataques, etc.) e essa avaliação pode ser obtida de maneira confiável através de uma auditoria por meio da corrente de blocos.

Capítulo 5

Trabalhos Relacionados

Alguns estudos vêm sendo desenvolvidos para investigação do problema de ataques em redes móveis via SS7. K. Jensen *et al.* desenvolvem um simulador próprio, onde são reproduzidos um cenário de uma rede de operadora bem como os ataques SS7 direcionados à mesma [51]. Após a geração das mensagens, é feita uma proposta de detecção de anomalias na rede SS7 através do algoritmo S-H-ESD (*Seasonal Hybrid Extreme Studentized Deviate*), que permite identificar a ocorrência dos ataques de interceptação SMS. Através de um conjunto de dados gerados pelo simulador, de 59.682 amostras, 6.495 foram detectadas como anomalias e estes resultados demonstraram a factibilidade do sistema proposto na detecção dos desvios [51]. Essa abordagem porém, não analisa um traço de dados reais, realizando a caracterização a partir de ataques gerados pelos autores.

Rupprecht *et al.* [23] apresentam um *survey* com as características das defesas adotadas contra os ataques em SS7 a partir de uma subdivisão baseada nas principais motivações dos atacantes. Os autores também desenvolvem uma metodologia para classificar os ataques por suas causas-raízes, das quais podem-se citar as deficiências nas especificações dos protocolos SS7 e na sua implementação, que é feita de forma insegura, com ausência de autenticação mútua e criptografia fraca. Para cada causa-raiz, é feita uma análise do impacto de cada ataque bem como a eficácia das propostas de defesa. Dentre as contramedidas citadas estão, por exemplo, o uso de *firewalls* que bloqueiam as principais mensagens de sinalização utilizadas em ataques já conhecidos; o uso de aplicativos para *Android* para detecção de *"IMSI Catchers"* (dispositivos MiM (*"Man-In-The-Middle"*) usados por atacantes para simular um elemento móvel e enviar mensagens legítimas para a rede das operadoras [52]); e o uso de sistemas de detecção de estações rádio base falsas (as quais captam assinantes da rede legítima [53]). Também foi proposto o uso do protocolo SSL (*Secure Sockets Layer*) para suprir a deficiência de autenticação mútua e de privacidade da comunicação em redes móveis [54].

Outro trabalho com uma abordagem parecida é o de Holtmanns *et al.* [55] onde

os autores também descrevem os tipos de ataques em SS7 assim como suas contramedidas, porém o foco do estudo é mostrar que mesmo as redes móveis mais recentes, como o 4G, também estão susceptíveis às mesmas vulnerabilidades que redes mais legadas (3G/2G), em situações de *roaming*. Os motivos são diversos, tais como a ausência de criptografia entre a comunicação do SS7 com o protocolo Diameter, que poderia ser provida por meio de protocolos como o IPSec. Além disso, os autores apontam que não é feita nenhuma filtragem de endereçamento IP na rede e nem alguma verificação de sanidade da rede. Por fim, a interconexão entre camadas em alguns casos não é feita diretamente com a operadora parceira para o *roaming* e sim por fornecedores, o que expõe ambas as redes pela confiança em uma rede de terceiros.

Dabrowski *et al.* apontam vários itens que podem ser utilizados como referência para definir se existe um dispositivo farejador na rede de acesso, como por exemplo, *Cell ID* incomum ou uma frequência de uso não usual, criptografia desabilitada e capacidades anormais de uso da célula [52]. Ademais, são desenvolvidos dois detectores de elementos de rede falsos, chamado "*IMSI Catcher Catcher*"ou ICC. O sICC, é um *hardware* desenvolvido com Raspberry Pi com acesso à rede 3G, apto a operar na faixa de 900 a 1800MHz. O mICC é um aplicativo em Android, com o mesmo propósito. Ambos foram efetivos na captura de características das células vizinhas (o sICC conseguiu capturas num range de 90km de distância) e se mostraram viáveis na captura de atacantes.

A literatura também apresenta a utilização de técnicas de aprendizado de máquina para detecção de anomalias na rede SS7 [51], permitindo identificar a ocorrência de um ataque. Em resumo, o autor desenvolve um simulador que reproduz uma rede de operadora com a geração de ataques em SS7, chamado *"SS7 Attack Simulator"*. Usando o algoritmo S-H-ESD (*Seasonal Hybrid Extreme Studentized Deviate*), os recursos propostos foram testados como uma indicação da viabilidade da aprendizagem de máquina na detecção do ataque de interceptação SMS. Através de um conjunto de dados de amostras geradas por meio do simulador de ataques SS7, um número de 6.495 amostras foram detectadas como anomalias e estes resultados demonstraram a factibilidade do sistema proposto na detecção dos desvios. Ao contrário deste, um conjunto de dados de tráfego real em SS7 foi analisado e discutido para verificação da vulnerabilidade à ataques em SS7 de uma grande operadora de telecomunicações que opera no Brasil.

Os benefícios do uso da corrente de blocos tais como transparência, imutabilidade e privacidade têm despertado interesse na resolução de problemas em redes móveis. Em trabalhos anteriores, Mafakheri *et al.* propuseram o uso da corrente de blocos para prover uma forma segura de autenticação e armazenamento de informações de usuários em redes 4G, através da descentralização do banco de dados

HSS (*Home Subscriber Server*) [15]. Na proposta, os processos de registro (*attach*) e desativação de usuários (*detach*) são feitos através de um contrato inteligente via corrente de blocos ao invés do procedimento normal utilizando o HSS. No entanto, o esquema proposto concentra-se na possibilidade de falha ou vulnerabilidade apenas do banco de dados da rede, não se preocupando com requisições legítimas originadas por usuários maliciosos que dispariariam ataques ainda que em um ambiente descentralizado.

A empresa Delloite apresentou um relatório mostrando alguns casos de uso onde a corrente de blocos pode ser útil aplicada a indústria de telecomunicações [14]. O primeiro caso a ser destacado é a utilização da tecnologia para prevenção de ataques de fraudes. Conforme já mencionado no Capítulo 2, caso uma operadora queira oferecer serviços de *roaming* em uma área fora de sua cobertura, a mesma precisa estabelecer parcerias através de contratos para interconexão com outras redes. Além disso, algumas operadoras funcionam como mediadoras nesta interconexão. Para o *roaming*, assim que um assinante sai da sua área local e se desloca para uma área visitada, a área visitada consulta o HLR da área local a fim de conhecer quais serviços podem ser disponibilizados para o usuário em situação de *roaming*. Feito isso, assim que o usuário começar a utilizar os serviços, os débitos são encaminhados para a rede de origem. Esse encaminhamento é feito por meio de uma operadora intermediária. Caso haja uma fraude nesta comunicação entre as operadoras, a provedora de serviços de origem não consegue tarifar o assinante, porém é cobrada pela parceria de *roaming*. Uma das razões para ocorrência das fraudes é devido ao tempo de resposta necessário para identificar uma fraude, que acaba sendo longo devido ao atraso no canal de comunicação entre as operadoras. Portanto, para minimizar este tempo, é proposto o uso de corrente de blocos privadas e permissionadas, onde o acordo de *roaming* é feito através de contratos inteligentes e as informações de faturamento são inseridas em transações. As principais contribuições são: a eliminação de uma terceira parte entre as operadoras, a redução das fraudes pela rapidez na mitigação de ataques e a disponibilização de um repositório de todas as transações aprovadas entre as partes. A segunda proposta da Delloite é para mitigar os ataques provenientes da ausência de segurança da rede, ou seja, aqueles onde o atacante se aproveita da exposição das informações sensíveis do usuário que são trafegadas por meio do canal de comunicação não-seguro. Quando um usuário liga o aparelho móvel, uma troca de mensagens de autenticação entre o equipamento e a estação base é feita, a fim de obter serviços. Como esta comunicação não é criptografada e o IMSI ou TMSI são divulgados por *broadcast* para todas as BTS's vizinhas, essas informações podem ser facilmente capturadas e posteriormente utilizadas na execução de um ataque. Babu *et al.* propuseram o uso da corrente de blocos como meio de inserção de segurança na rede [14]. Ao invés da divulgação por meio de *broadcast*,

os elementos seriam inseridos na corrente de blocos e utilizariam criptografia com chaves privadas e públicas para a comunicação, onde somente a chave pública é divulgada. Ademais, como todas as estações base são nós pertencentes a uma corrente de blocos comum, não há necessidade de autenticação do usuário cada vez que este mude de BTS, pois tais informações públicas já são conhecidas por todos os nós através da replicação das informações dentro da cadeia de blocos.

Capítulo 6

Conclusões e Trabalhos Futuros

O Sistema de Sinalização por Canal Comum nº7 foi considerado uma disruptura para a sua época, pois a partir deste sistema, os canais de sinalização passaram a ter um circuito dedicado para tráfego de informações, aumentando a capacidade dos canais de voz. Com o passar dos anos, o sistema foi atualizado para incorporar novas tecnologias, como o IP (através do SIGTRAN) e as redes móveis através do conjunto de protocolos MAP. O SS7 é utilizado até hoje, para interconexão de redes e comunicação entre os elementos das redes móveis, 2G e 3G. Originalmente, o SS7 baseava-se em relações de mútua confiança entre as operadoras, já que estas eram poucas e geralmente ligadas ao governo. Essa relação de confiança foi quebrada a partir do crescimento do número de operadoras e do acesso à Internet, trazendo à tona problemas de segurança que antes não existiam. Os atacantes podem explorar as vulnerabilidades do SS7 como os problemas advindos da falta de segurança na especificação do protocolo [23] e gerar vários tipos de ataques de forma combinada. Apesar da evolução tecnológica, mesmo as gerações mais modernas como o 4G permanecem à sombra da insegurança do SS7, por não possuir ainda a totalidade da rede com o protocolo Diameter e por não permitir o serviço de voz nativamente em sua tecnologia. O primeiro passo para qualquer iniciativa de mitigação de ataques é a investigação e leitura do cenário de rede, a fim de promover uma auditoria dos ataques encontrados. A partir desse conhecimento, é possível elaborar formas mais efetivas de combate aos problemas encontrados, detectando os ataques e tratando caso a caso, já que para uma solução definitiva, o ideal seria a completa substituição dos protocolos, o que não deve acontecer tão cedo. De qualquer forma, caracterizar vulnerabilidades e ameaças é sempre importante antes de propor novas soluções de segurança em redes.

As principais contribuições deste trabalho são: a caracterização da vulnerabilidade da rede de uma grande provedora de telecomunicações através da coleta de dados reais; a inserção da tecnologia da corrente de blocos (*Blockchain*) como forma de auditoria para o problema de ataques SS7 em redes móveis e a avaliação da pro-

posta como contrato inteligente e de desempenho através de ferramentas conhecidas e disponíveis.

A análise do presente trabalho é revelante e atual, sobretudo após o incidente de segurança envolvendo o vazamento de conversas da Operação Lava Jato , que mostrou os ataques de interceptação e rastreamento sendo executados. Os principais resultados mostraram a grande quantidade de ameaças disparadas contra a rede da operadora (no total, as tentativas somaram 3.285.005 ameaças no período observado). Além disso, a análise das ameaças permitiu a descoberta de aspectos antes desconhecidos, como a grande frequência das ameaças, que mostra o quão exposta está a rede da operadora de telecomunicações. Estes dados servem para afirmar que o provedor de serviços deve tomar as providências de segurança para as vulnerabilidades em SS7 com urgência a fim de minimizar os riscos que a rede está submetida e para proteção de dados dos clientes.

Ademais, foi feita a proposta de auditoria através da tecnologia de correntes de blocos. A análise de desempenho feita no *Hyperledger Caliper* a partir do código gerado permitiu verificar a adoção da solução em um ambiente real. A vazão máxima encontrada de 61,3tps mostra que o modelo de corrente de blocos pode ser implementado com algoritmos de consenso que utilizem poucos nós por possuírem maior escalabidade [48, 49]. O consumo dos recursos de memória e CPU para o processamento das transações mostrou que é possível implantar um sistema baseado em correntes de blocos no *hardware* utilizado atualmente nas operadoras, que sobrepõe a máquina virtual utilizada na execução dos testes de desempenho. Ou seja, existe a possibilidade das operadoras de telecomunicações adotarem a corrente de blocos como forma de melhoria na segurança da rede sem grandes impactos no cenário atual de recursos computacionais.

Como trabalho futuro, será feita uma revisão do cenário da proposta, estendendo a mesma para um ambiente distribuído de multi-operadoras. O objetivo é aplicar um algoritmo de consenso que execute o modelo de corrente de blocos proposto entre as operadoras, em situação de *roaming* e vulnerabilidades em SS7, conferindo segurança às informações trocadas entre as redes. Além disso, objetiva-se a extensão da proposta com o uso de outros algoritmos de consenso, como o PBFT, pois em um cenário híbrido entre operadoras concorrentes não há confiança mútua entre as partes (exceto em casos de contratos fixados). Também pretende-se utilizar a plataforma *Ethereum* ao invés do *Hyperledger Fabric* e verificar se a escolha pelo HLF foi a melhor a ser adotada. Ademais, pretende-se alterar o dimensionamento da máquina virtual adotada para verificar o impacto das configurações da mesma nos resultados encontrados.

Adicionalmente, outro objetivo é realizar a caracterização e identificação de vulnerabilidades em outros tipos de protocolos de sinalização, como o Diameter e o GTP

(*GPRS Tunneling Protocol*). O GTP é baseado no protocolo IP, executado nos roteadores que interligam as redes móveis e responsável pelo encapsulamento de dados GPRS entre os elementos da móvel. Além disso, pretende-se realizar a modelagem dos atacantes de forma individual, ainda utilizando os dados desta dissertação.

Ademais, um estudo futuro interessante é o de detecção dos ataques a partir de mensagens em SS7, que podem ser geradas através de um simulador. Esta detecção eliminaria a necessidade de um intermediário como o fornecedor do *firewall* e possibilitaria a redução de custos a serem investidos pela operadora.

Apêndice A

Código em Go

```
package main

import (
        "fmt"
        "strconv"
        "encoding/base64"
        "crypto/sha256"

        "github.com/hyperledger/fabric/core/chaincode/shim"
        pb "github.com/hyperledger/fabric/protos/peer"
)

type SimpleChaincode struct {
}

type Transaction struct {
        Author       string
}

type argMAP struct {
    UserA  string //user
    cgPA   string //callingPartyAddress
}

//==================MAIN=================

func main() {
        err := shim.Start(new(SimpleChaincode))
        if err != nil {
                fmt.Printf("Error starting Chaincode: %s", err)
        }
}

```

```
34 //==================INIT=================
35
36 func (t *SimpleChaincode) Init(stub shim.ChaincodeStubInterface) pb
       .Response {
37         return shim.Success(nil)
38 }
39
40 //==================INVOKE=================
41
42 func (t *SimpleChaincode) Invoke(stub shim.ChaincodeStubInterface)
       pb.Response {
43         function, args := stub.GetFunctionAndParameters()
44         if function == "setTransaction" {
45                 return t.setTransaction(stub, args)
46         } else if function == "setMAP" {
47                 return t.setMAP(stub, args)
48         } else if function == "delete" {
49                 return t.delete(stub, args)
50         } else if function == "query" {
51                 return t.query(stub, args)
52         }
53
54         return shim.Error("Invalid invoke function name. Expecting
       \"setTransaction\" \"setMAP\" \"delete\" \"query\"")
55 }
56
57 //==================FUNCTIONS=================
58
59 //Set the transaction arguments
60 func (t *SimpleChaincode) setTransaction(stub shim.
       ChaincodeStubInterface, args []string) pb.Response {
61         var err error
62
63         if len(args) != 1 {
64                 return shim.Error("Incorrect number of arguments.
       Expecting 1.\n Usage: `{\"Args\":[\"[Author]\"]}`")
65         }
66
67         if len(args[0]) <= 0 {
68                 return shim.Error("Argument must be a non-empty
       string.")
69         }
70
71         // Initialize the chaincode
72         Author := args[0]
73         if err != nil {
74                 return shim.Error("Expecting string value for
```

```
    Author.")
75          }
76
77          // calculate the hash value as identity
78          hash := sha256.Sum256([]byte(Author))
79          sessionId := base64.URLEncoding.EncodeToString(hash[:])
80
81          //Time stamp
82          CreateTime, err := stub.GetTxTimestamp()
83          if err != nil {
84              return shim.Error("Expecting string value for Author.")
85          }
86
87          fmt.Printf("Author = %d, sessionId = %d\n, CreateTime = %d\
    n", Author, sessionId, CreateTime)
88
89          // check if sessionId already exists
90          state, err := stub.GetState(sessionId)
91          if state != nil && err == nil {
92                  return shim.Error("Identity already published");
93          }
94
95          // Write the state to the ledger
96          err = stub.PutState(Author, []byte(Author))
97          if err != nil {
98                  return shim.Error(err.Error())
99          }
100         err = stub.PutState(sessionId, []byte(sessionId))
101         if err != nil {
102                 return shim.Error(err.Error())
103         }
104         //Return sucess
105         return shim.Success(nil)
106
107 }
108
109 //Set the user arguments
110 func (t *SimpleChaincode) setMAP(stub shim.ChaincodeStubInterface,
    args []string) pb.Response {
111         // similar to "InsertSubscriberData"
112
113         var err error
114
115         if len(args) != 2 {
116                 return shim.Error("Incorrect number of arguments.
    Expecting 4.\n Usage: `{\"Args\":[\"[UserA]\",\"[
    callingPartyAddress with 13 Digits]\"]}`")
```

```
117         }
118
119         if len(args[0]) <= 0 { //If the argument is empty, the
    trasaction is discarded
120         return shim.Error("Argument must be a non-empty string.")
121         }
122
123         if len(args[1]) <= 0 {
124         return shim.Error("Argument must be a non-empty string.")
125         }
126
127         if len(args[1]) != 13 { //13 Digits is the length defined
    by 3GPP for callingPartyAddress and/or calledPartyAddress
128         return shim.Error("Argument must be a value with 13 digits
    .")
129         }
130
131         // Initialize the chaincode
132         UserA := args[0]
133         Userbytes,err := stub.GetState(UserA)
134          if Userbytes != nil {
135                 return shim.Error("User already exists.")
136         }
137
138         cgPA,err := strconv.Atoi(args[1])
139         if err != nil {
140                 return shim.Error("Expecting integer value for
    callingPartyAdress.")
141         }
142
143         fmt.Printf("User = %d\n, callingPartyAdress = %d\n", UserA,
     cgPA)
144
145         // Write the state to the ledger
146         err = stub.PutState(UserA, []byte(strconv.Itoa(cgPA)))
147         if err != nil {
148                 return shim.Error(err.Error())
149         }
150
151         return shim.Success(nil)
152 }
153
154 // Deletes an user from database
155 func (t *SimpleChaincode) delete(stub shim.ChaincodeStubInterface,
    args []string) pb.Response {
156         if len(args) != 1 {
157                 return shim.Error("Incorrect number of arguments.
```

```
    Expecting 1.")
158         }
159
160         UserA := args[0]
161
162         // Delete the User from the state in ledger
163         err := stub.DelState(UserA)
164         if err != nil {
165                 return shim.Error("Failed to delete user!")
166         }
167
168         return shim.Success(nil)
169 }
170
171 //Query an user
172 func (t *SimpleChaincode) query(stub shim.ChaincodeStubInterface,
    args []string) pb.Response {
173         //similar to sendAuthenticationInfo or sendRoutingInfo
174
175         if len(args) != 1 {
176                 return shim.Error("Incorrect number of arguments.
    Expecting name of the user to query.")
177         }
178
179         // Get the state from the ledger
180         Userbytes, err := stub.GetState(args[0])
181         if err != nil {
182                 return shim.Error("Error, not found!")
183         } else if Userbytes == nil {
184                 return shim.Error("Error, not found!")
185         }
186         return shim.Success(Userbytes)
187 }
```

Listing A.1: Contrato inteligente (*Chaincode*) em Go.

Referências Bibliográficas

[1] SRIVASTAVA, L. "Mobile phones and the evolution of social behaviour", *Behaviour & information technology*, v. 24, n. 2, pp. 111–129, 2005.

[2] ATZORI, L., IERA, A., MORABITO, G. "The internet of things: A survey", *Computer networks*, v. 54, n. 15, pp. 2787–2805, 2010.

[3] FETTWEIS, G. P. "The tactile internet: Applications and challenges", *IEEE Vehicular Technology Magazine*, v. 9, n. 1, pp. 64–70, 2014.

[4] INTELLIGENCE, G. *The Mobile Economy 2019*. Relatório técnico, 2019. Disponível em: <https://www.gsma.com/r/mobileeconomy/>.

[5] DRYBURGH, L., HEWETT, J. *Signaling System No. 7 (SS7/C7): Protocol, Architecture, and Services (Networking Technology)*. Cisco Press, 2004.

[6] ENGEL, T. "Locating Mobile Phones using Signalling System #7". 25th Chaos Communication Congress, 2008. Disponível em: <https://berlin.ccc.de/~tobias/25c3-locating-mobile-phones.pdf>. [Online; Acessado em 06/10/2017].

[7] *Official Document IR.70 - SMS SS7 Fraud*. Relatório técnico, GSM Association, 2013. Disponível em: <https://www.gsma.com/newsroom/wp-content/uploads/IR.70-v4.0.pdf>.

[8] *Official Document IR.77 - Inter-Operator IP Backbone Security Requirements For Service Providers and Inter-operator IP backbone Providers*. Relatório técnico, GSM Association, 2007. Disponível em: <https://www.gsma.com/publicpolicy/wp-content/uploads/2012/03/ir77.pdf>.

[9] *Official Document IR.88 - LTE and EPC Roaming Guidelines*. Relatório técnico, GSM Association, 2016. Disponível em: <https://www.gsma.com/newsroom/wp-content/uploads/IR.88-v15.0.pdf>.

[10] RAO, S. P., KOTTE, B. T., HOLTMANNS, S. "Privacy in LTE Networks". In: *9th EAI International Conference on Mobile Multimedia Communications*, pp. 176–183. ACM, 2016.

[11] TU, G.-H., LI, C.-Y., PENG, C., et al. "How voice call technology poses security threats in 4G LTE networks". In: *2015 IEEE Conference on Communications and Network Security (CNS)*, pp. 442–450. IEEE, 2015.

[12] ROTH, J. D., TUMMALA, M., MCEACHEN, J. C., et al. "On location privacy in LTE networks", *IEEE Transactions on Information Forensics and Security*, v. 12, n. 6, pp. 1358–1368, 2017.

[13] BAUTISTA, J. E. V., SAWHNEY, S., SHUKAIR, M., et al. "Performance of CS Fallback from LTE to UMTS", *IEEE Communications Magazine*, v. 51, n. 9, pp. 136–143, 2013.

[14] BABU, A., DAVIS, B., BRUWER, T., et al. "How Blockchain can impact the telecommunications industry", 2018.

[15] MAFAKHERI, B., SUBRAMANYA, T., GORATTI, L., et al. "Blockchain-based Infrastructure Sharing in 5G Small Cell Networks". In: *14th International Conference on Network e Service Management (CNSM)*, pp. 313–317, dez. 2018.

[16] NASIR, Q., QASSE, I. A., ABU TALIB, M., et al. "Performance analysis of hyperledger fabric platforms", *Security and Communication Networks*, v. 2018, 2018.

[17] SUKHWANI, H., WANG, N., TRIVEDI, K. S., et al. "Performance Modeling of Hyperledger Fabric (Permissioned Blockchain Network)". In: *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, pp. 1–8. IEEE, 2018.

[18] GORENFLO, C., LEE, S., GOLAB, L., et al. "Fastfabric: Scaling hyperledger fabric to 20,000 transactions per second", *arXiv preprint arXiv:1901.00910*, 2019.

[19] ONG, L., RYTINA, I., GARCIA, M., et al. *Framework Architecture for Signaling Transport (RFC 2719)*. Relatório técnico, IETF, out. 1999. Disponível em: <`https://tools.ietf.org/html/rfc2719`>.

[20] 3GPP. *Mobile Application Part (MAP) specification*. Technical specification (ts), 3rd Generation Partnership Project (3GPP), 1999. Disponível em: <`https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1585`>.

[21] TECHNOLOGIES, P. "SS7 Vulnerabilities e Attack Exposure Report". 2018. Disponível em: <`https://bit.ly/2JgBhYl`>.

[22] NOHL, K. *Mobile Self-Defense*. Relatório técnico, 2014. Disponível em: <https://events.ccc.de/congress/2014/Fahrplan/system/attachments/2493/original/MobileSelfDefense-KarstenNohl-31C3-v1.pdf>.

[23] RUPPRECHT, D., DABROWSKI, A., HOLZ, T., et al. "On security research towards future mobile network generations", *IEEE Communications Surveys & Tutorials*, v. 20, n. 3, pp. 2518–2542, 2018.

[24] KALENDERI, M., PNEVMATIKATOS, D., PAPAEFSTATHIOU, I., et al. "Breaking the GSM A5/1 cryptography algorithm with rainbow tables and high-end FPGAS". pp. 747–753, 08 2012. ISBN: 978-1-4673-2257-7. doi: 10.1109/FPL.2012.6339146.

[25] MEYER, U., WETZEL, S. "On the impact of GSM encryption and man-in-the-middle attacks on the security of interoperating GSM/UMTS networks". In: *2004 IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE Cat. No. 04TH8754)*, v. 4, pp. 2876–2883. IEEE, 2004.

[26] GENDRULLIS, T., NOVOTNỲ, M., RUPP, A. "A real-world attack breaking A5/1 within hours". In: *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 266–282. Springer, 2008.

[27] BAIS, A., PENZHORN, W. T., PALENSKY, P. "Evaluation of UMTS security architecture and services". In: *2006 4th IEEE International Conference on Industrial Informatics*, pp. 570–575. IEEE, 2006.

[28] GOLD, S. "Cracking cellular networks via femtocells", *Network Security*, v. 2011, n. 9, pp. 5–8, 2011.

[29] AVIZIENIS, A., LAPRIE, J.-C., RANDELL, B., et al. "Basic concepts and taxonomy of dependable and secure computing", *IEEE transactions on dependable and secure computing*, v. 1, n. 1, pp. 11–33, 2004.

[30] *SS7 Security Network Implementation Guidelines*. Relatório técnico, GSM Association, 2016. Version 5.0.

[31] VAN DEN BROEK, F., VERDULT, R., DE RUITER, J. "Defeating IMSI catchers". In: *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 340–351. ACM, 2015.

[32] CORLETTI, A. "Análisis de ataques/vulnerabilidades SS7/Sigtran empleando Wireshark (y/o tshark) y Snort". `https://bit.ly/2p37dWI`, 2018. [Online; Acessado em 29/07/2019].

[33] NAKAMOTO, S. *Bitcoin: A peer-to-peer electronic cash system.* Relatório técnico, Bitcoin, 2008. Disponível em: <`https://bitcoin.org/bitcoin.pdf`>.

[34] DORRI, A., KANHERE, S. S., JURDAK, R., et al. "Blockchain for IoT security and privacy: The case study of a smart home". In: *2017 IEEE international conference on pervasive computing and communications workshops (PerCom workshops)*, pp. 618–623. IEEE, 2017.

[35] REYNA, A., MARTÍN, C., CHEN, J., et al. "On blockchain and its integration with IoT. Challenges and opportunities", *Future Generation Computer Systems*, v. 88, pp. 173–190, 2018.

[36] NOVO, O. "Blockchain meets IoT: An architecture for scalable access management in IoT", *IEEE Internet of Things Journal*, v. 5, n. 2, pp. 1184–1195, 2018.

[37] GORDON, W. J., CATALINI, C. "Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability", *Computational and structural biotechnology journal*, v. 16, pp. 224–230, 2018.

[38] METTLER, M. "Blockchain technology in healthcare: The revolution starts here". In: *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–3. IEEE, 2016.

[39] LIANG, X., ZHAO, J., SHETTY, S., et al. "Integrating blockchain for data sharing and collaboration in mobile healthcare applications". In: *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5. IEEE, 2017.

[40] LIN, J., SHEN, Z., ZHANG, A., et al. "Blockchain and iot based food traceability for smart agriculture". In: *Proceedings of the 3rd International Conference on Crowd Science and Engineering*, p. 3. ACM, 2018.

[41] CASADO-VARA, R., PRIETO, J., DE LA PRIETA, F., et al. "How blockchain improves the supply chain: Case study alimentary supply chain", *Procedia computer science*, v. 134, pp. 393–398, 2018.

[42] WÜST, KARL E GERVAIS, A. "Do you need a Blockchain?" In: *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pp. 45–54. IEE, 2018.

[43] ZHENG, Z., XIE, S., DAI, H.-N., et al. "Blockchain challenges and opportunities: A survey", *International Journal of Web and Grid Services*, v. 14, n. 4, pp. 352–375, 2018.

[44] DE ANGELIS, S., ANIELLO, L., BALDONI, R., et al. "PBFT vs Proof-of-Authority: applying the CAP theorem to permissioned blockchain". In: *Italian Conference on Cyber Security (ItaSec)*, pp. 1–11, jan. 2018.

[45] ETHEREUM. *Parity: next generation Ethereum browser.* 2015. Disponível em: <`https://www.parity.io/`>.

[46] DINH, T. T. A., WANG, J., CHEN, G., et al. "BLOCKBENCH: A framework for analyzing private blockchains". In: *ACM International Conference on Management of Data (SIGMOD)*, pp. 1085–1100, maio 2017.

[47] BACH, L. M., MIHALJEVIĆ, B., ŽAGAR, M. "Comparative analysis of blockchain consensus algorithms". In: *41st International Convention on Information e Communication Technology, Electronics e Microelectronics (MIPRO)*, pp. 1545–1550, maio 2018.

[48] LI, C., LI, P., ZHOU, D., et al. "Scaling nakamoto consensus to thousands of transactions per second", *arXiv preprint arXiv:1805.03870*, 2018.

[49] VUKOLIĆ, M. "The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication". In: *International Workshop on Open Problems in Network Security (iNetSec)*, pp. 112–125, out. 2015.

[50] ANDROULAKI, E., BARGER, A., BORTNIKOV, V., et al. "Hyperledger fabric: a distributed operating system for permissioned blockchains". In: *Proceedings of the Thirteenth EuroSys Conference*, p. 30. ACM, 2018.

[51] JENSEN, K., DO, T. V., NGUYEN, H. T., et al. "Better Protection of SS7 Networks With Machine Learning". In: *6th International Conference on IT Convergence e Security (ICITCS)*, pp. 1–7, set. 2016.

[52] DABROWSKI, A., PIANTA, N., KLEPP, T., et al. "IMSI-Catch Me If You Can: IMSI-Catcher-Catchers". In: *30th Annual Computer Security Applications Conference (ACSAC)*, pp. 246–255, dez. 2014.

[53] LI, Z., WANG, W., WILSON, C., et al. "FBS-Radar: Uncovering Fake Base Stations at Scale in the Wild". In: *Network and Distributed System Security Symposium (NDSS)*, pp. 1–15, fev. 2017.

[54] KAMBOURAKIS, G., ROUSKAS, A., GRITZALIS, S. "Performance evaluation of public key-based authentication in future mobile communication systems", *EURASIP Journal on wireless Communications e Networking*, v. 2004, n. 1, pp. 184–197, 2004.

[55] HOLTMANNS, S., RAO, S. P., OLIVER, I. "User location tracking attacks for LTE networks using the interworking functionality". In: *2016 IFIP Networking conference (IFIP Networking) and workshops*, pp. 315–322. IEEE, 2016.